



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶: C12N 15/86, 7/00, 15/62, A61K 39/145, 35/76	A1	(11) International Publication Number: WO 96/10641 (43) International Publication Date: 11 April 1996 (11.04.96)
(21) International Application Number: PCT/EP95/03663 (22) International Filing Date: 18 September 1995 (18.09.95) (30) Priority Data: 94115505.3 30 September 1994 (30.09.94) EP (34) Countries for which the regional or international application was filed: DE et al. (71) Applicant (for all designated States except US): BAYER AKTIENGESELLSCHAFT [DE/DE]; D-51368 Leverkusen (DE). (72) Inventors; and (75) Inventors/Applicants (for US only): HOBOM, Gerd [DE/DE]; Arndtstrasse 14, D-35392 Giessen (DE); NEUMANN, Gabriele [DE/DE]; Hermann-Löns-Strasse 7, D-63447 Maintal (DE); MENKE, Annette [DE/DE]; Ockershäuser Allee 7, D-35037 Marburg (DE). (74) Common Representative: BAYER AKTIENGESELLSCHAFT; D-51368 Leverkusen (DE).		(81) Designated States: AU, CA, CN, FI, JP, KR, MX, NO, NZ, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
(54) Title: AN ATTENUATED VACCINATION AND GENE-TRANSFER VIRUS, A METHOD TO MAKE THE VIRUS AND A PHARMACEUTICAL COMPOSITION COMPRISING THE VIRUS		
(57) Abstract <p>RNA polymerase I transcription <i>in vivo</i> in transiently DNA-transfected cells has been used for expression of influenza vRNA molecules coding for chloramphenicol acetyltransferase (CAT) in anti-sense orientation. Influenza virus superinfection served to provide viral RNA polymerase and other proteins for transcriptional conversion of minus-strand vRNA into plus-strand viral mRNA molecules expressing CAT activity. This system has been used for an analysis via nucleotide exchanges as well as deletions and insertions of both terminal segments of the vRNA sequence which cooperatively constitute the vRNA promoter structure. Several mutants with greatly enhanced expression rates over wild-type levels have been constructed, which also can be packaged and serially passaged into progeny virus. The data obtained for the mutations in various promoter elements support a model of consecutive, doublestrand vRNA promoter structures in binding of viral polymerase and initiation of RNA synthesis. Preparations of attenuated influenza virus for vaccination purposes include a single recombinant segment with promoter up mutation(s) for over-expression of an own or foreign gene product, which at the same time because of its over-replication serves to decrease the number of helper virus RNP segments. The same viruses further have been passaged through a step of ribozyme cleavage acting at one of the helper viral segments, which will delete this vital function and structure with high rates from the virus progeny. The resulting attenuated viruses will interact with their target cells in only one round of abortive infection, and are unable to produce viral progeny.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LV	Latvia	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

**AN ATTENUATED VACCINATION AND GENE-TRANSFER VIRUS, A
METHOD TO MAKE THE VIRUS AND A PHARMACEUTICAL COMPO-
SITION COMPRISING THE VIRUS**

The object of the present invention was to make a vaccination virus. This objective has been fulfilled with the segmented virus constructed as described herein.

The genome of influenza A viruses consists of 8 different single-stranded viral RNA (vRNA) molecules of negative polarity, which have in common 5' and 3' terminal sequences largely complementary to each other. These conserved segments 13 and 12 nucleotides in length are known to form double-stranded RNA panhandle structures (Hsu et al., 1987; Fodor et al., 1993) which have been analysed in more detail recently *in vitro* using internally deleted model RNAs (Baudin et al., 1994; Tiley et al., 1994). In the virion the panhandle ends of all RNA segments are found in specific binding to viral RNA polymerase complexes, while the remaining internal segments stay single-stranded with viral nucleoprotein (NP) in cooperative binding (Compans et al., 1972; Honda et al., 1988; Martin et al., 1992). Upon infection these viral RNPs initially serve as templates for the synthesis of viral mRNAs by a specific cap-snatching mechanism (Plotch et al., 1979; Braam et al., 1983), and later on will direct synthesis of full-length complementary RNAs (cRNAs), probably dependent on the absence or presence of newly synthesized NP protein (Shapiro and Krug, 1988). The plus-strand cRNAs are then used as templates for progeny vRNA synthesis.

The viral RNA polymerase complex consisting of proteins PB1, PB2, and PA is involved in all three different modes of RNA synthesis during the viral replication cycle, following its specific binding to the terminal panhandle segments of both vRNAs and cRNAs. Sequence comparison reveals that the vRNA and cRNA

termini have similar, but not identical sequences. For that reason vRNA and cRNA recognition may be distinguished because of these structural alterations allowing for asymmetries in initiation of plus and minus strand RNA synthesis, and possibly in viral RNP packaging, which has also been suggested to be controlled by the panhandle RNA sequence (Hsu et al., 1987).

Recently, we reported on an *in vivo* system for the introduction of specific mutations into the genome of influenza viruses: viral cDNA has been inserted in antisense orientation between mouse rDNA promoter and terminator sequences. This has been derived from *in vitro* transcription experiments based on nuclear extracts from Ehrlich ascites cells, which resulted in transcripts exactly resembling influenza vRNA. For a series of *in vivo* studies, the viral coding sequence was replaced by the coding sequence for chloramphenicol-acetyltransferase (CAT), however, with both influenza terminal non-coding sequences being retained exactly on the resulting vRNA transcripts. After transfection of this recombinant DNA template into mouse cells followed by influenza virus infection, CAT activity was detectable. Transfer of supernatants to different cells demonstrated that CAT-vRNAs transcribed *in vivo* by cellular RNA polymerase I were not only transcribed by viral RNA polymerase into plus-strand mRNA and translated into CAT protein, but also were replicated and packaged into infectious progeny virus particles (Zobel et al., 1993; Neumann et al., 1994).

We have used this system for a stepwise introduction of single and multiple mutations into the conserved panhandle RNA sequence, thereby effectively converting the HA-vRNA promoter sequence into an HA-cRNA promoter sequence and vice versa. For these series of constructs CAT activities have been measured both in primarily transfected and infected B82 cells and, after passaging of B82 supernatants, in secondarily infected MDCK cells. From the results obtained we propose a model for the terminal RNA sequence as being recognized RNA polymerase in consecutive steps of different structure when used as a template for initiation of viral mRNA synthesis.

The present invention relates to a segmented RNA virus which comprises one or more segments which have been genetically modified to show improved transcription, replication and/or expression rates.

The virus can be one where one or more modifications have been introduced in the noncoding region(s) and/or one or more modifications have been introduced in the coding region(s). A possibility is that at least one modified segment is derived from an original one by sequence variation(s). It is also possible that at least one modified segment is an artificial addition to the set of original or modified original segments. The virus can be one wherein the modified segment comprises a nucleotide sequence which codes for a protein or peptide which is foreign to the original virus. Preferred is that the foreign protein or peptide constitutes an antigen or antigen-like sequence, a T-cell epitope or related sequence. In such a case it is possible that the segment comprises repetitions of an antigen or epitope or other. Such an antigen or epitope can be for example derived from HIV, Herpes-Virus, Human Papilloma virus, Rhinovirus, CMV or Hog Cholera Virus (HCV). The virus of the present invention may be a single stranded negative-strand RNA virus as for example one of the Orthomyxoviridae family, the Bunyaviridae family or of the Arenaviridae family. The most preferred virus is an influenza virus. The virus of the present invention can also be a double-stranded RNA virus, as for example a reovirus, a rotavirus or an orbivirus. The viruses of the present invention can be used in gen-therapy.

The present invention also relates to the virus and use of the virus for the preparation of pharmaceuticals.

Mutational analysis of vRNA 3' terminal sequence positions.

Influenza A viral RNA 5' and 3' ends have similar, but not identical sequences with nucleotide mismatches at positions 3, 5 and 8, and an additional unpaired nucleotide is located at position 10 in the 5' region. Nevertheless, both vRNA termini hybridize into a double stranded panhandle structure made up of twelve and thirteen nucleotides in common for all eight RNA segments, plus in average three additional basepairs specific for each of the vRNA molecules. Due to the deviations mentioned the cRNA or plus strand panhandle structures have to be different from the vRNA structures; however, both are recognized by viral RNA polymerase and are used for initiation of RNA synthesis, i.e. they are constituting a promoter structure. Even if in initial recognition and binding of RNA polymerase the double stranded RNA panhandle structure is known to be the substrate, and is also observed in virion RNPs (Hsu et al., 1987), for the initiation step of transcription at the 3' ultimate template position this terminal region has to

be separated into a partially single stranded, i.e. 'forked' structure (Fodor et al., 1994). RNA polymerase may be predicted to continue its binding interaction with both, the remaining double stranded segment: nucleotides $\overline{10}$ to $\overline{15}$ versus 11 to 16¹, and to the single stranded 3' template segment: nucleotides $\overline{1}$ up to $\overline{9}$, as well as the 5' single stranded end (Tiley et al., 1994). Introduction of mutations at specific positions in either strand may hence alterate simultaneously both of these consecutive vRNA promoter structures: panhandle and fork in different ways, and will in addition also result in corresponding variations of the cRNA promoter structure.

10 To investigate the importance of the three mismatch positions, specific single, double or triple nucleotide exchanges were first introduced into the vRNA 3' end sequence at positions $\overline{3}$, $\overline{5}$ and $\overline{8}$, thereby approaching a fully double-stranded vRNA promoter structure, in a step-wise manner. At the same time the vRNA 3' end template sequence will become equivalent to the cRNA 3' end in these positions, but not in regard to the additional nucleotide at position 10. Single
15 nucleotide exchanges according to this scheme (pHL1098, pHL1099, pHL1100) abolished the promoter activity, and no CAT activity was observed, as has been reported before with a different method (Luo et al., 1993). Two of the double mutation constructs (pHL1101, pHL1103) also gave negative results.

20 In contrast, for pHL1102 (G3A, U8C)¹, a significant CAT activity was detected, distinctly higher than for the corresponding wild-type construct (pHL926;) which in the conditions applied (8 hr after infection) resulted in rather low levels of CAT expression. This activity increase is further enhanced for the final construct of this series carrying the triple exchange G $\overline{3}$ A, C5U and U $\overline{8}$ C (pHL1104), i. e. trans-
25 fection of pHL1104 DNA followed by influenza virus infection resulted in a very high level of CAT expression, also considerably above the pHL1102 results.

These results have been repeated using various conditions of transfection and infection as well as determining kinetic data during the course of infection. While the pHL1104 variant is always observed far superior over any wild-type construct

30 ¹Notations concerning nucleotides $\overline{1}$ to $\overline{15}$ refer to positions in the vRNA 3' end, e.g. position $\overline{2}$ designates the penultimate nucleotide; 5' end positions are given in ordinary numbers. The notation G $\overline{3}$ A describes a mutational change of guanosine to adenosine at position $\overline{3}$.

that expression ratio may be variable and difficult to quantitate (between around 20 fold and nearly 100 fold). Rather short infective cycles of eight hours as used prevalently appear to put more slowly replicating, i.e. wild-type molecules at a disadvantage, in particular in passaging of packaged pseudo-vRNA molecules via virus progeny, both is found increased for wildtype and related constructs after DNA transfection plus *twelve* hours of infection (see Neumann et al., 1994). Remaining deviations in CAT expression ratios may be attributed to variations in growth conditions in individual experiments.

Mutational analysis of vRNA 5' terminal sequence positions

We also addressed the question whether the unexpectedly high viral mRNA expression rate of pHL1104 is the consequence of a stabilized panhandle double-strand structure or may be directly attributed to the point mutations introduced into the vRNA 3' sequence, and active when being used as a single-stranded template segment, e.g. in the 'forked' structure.

For this purpose we constructed pHL1124, three complementary point mutations introduced at the 5' end of the vRNA sequence again in positions 3, 5 and 8 (U3C, G5A, A8G). Together with a sequence wild-type vRNA 3' end these variations again result in a panhandle structure free of mismatches and, therefore, pHL 1124 is equivalent in this regard to pHL1104, but different in the sequence of its template and non-template single strands. No significant CAT expression was detected for pHL1124. We conclude that the increased CAT activity of pHL1104 is not a consequence of the stabilized panhandle structure itself, but at least in part is a consequence of the individual nucleotide exchanges at positions $\overline{3}$, $\overline{5}$ and $\overline{8}$ at the 3' end of the vRNA sequence, it is also more likely then to originate from other structural intermediates of initiation than a stabilized panhandle.

Mutational analyses of concerted exchanges at both ends of the vRNA sequence

In order to determine in detail the influence of single, double and triple exchanges at the vRNA 5' end upon CAT expression rates we also used the improved vRNA 3' end sequences of pHL1104 and pHL1102 as starting points rather than the corresponding wild-type sequence. From the series of experiments related to pHL1104 and from the equivalent series related to pHL 1102 it can be concluded that retaining a G residue in position 5 is the most important single feature in

these 5' end variations. A single exchange into an A residue at position 5 as in pHL1185 will render the promoter entirely inactive, while single exchanges in positions 3 or 8, as well as a 3 plus 8 double exchange will retain promoter activity even if reduced from the level observed for pHL1104, but still above wild-type expression rates. While the G5A nucleotide substitution opposite nucleotide C $\bar{5}$ in the 3' terminus results in losing one basepair (in the panhandle context) within the pHL1104 series, a basepair is indeed gained by exactly the same G5A exchange within the pHL1102 series, i.e. opposite the U $\bar{5}$ residue as present in the pHL1102 vRNA 3' end. Since again the G5A exchange results in loss of promoter function inspite of gaining one basepair we conclude that the guanosine at position 5 may be important for RNA polymerase binding within the 5' non-template single strand rather than being part of the panhandle double-stranded structure in this region. The importance of a G residue at this position has been shown earlier in a single-step mutational analysis (Li and Palese, 1992), while non-template strand binding of RNA polymerase has been studied recently *in vitro* (Tiley et al., 1994). Different from the deleterious effect of an exchange at position 5 exchanges at positions 3 and in particular 8 are of minor importance. The series of 5' nucleotide exchanges has also been repeated for the pHL1102 version of the vRNA 3' end yielding exactly the same pattern of results, albeit at the somewhat reduced levels characteristic for pHL1102. The only result in both series not quite in agreement with a uniquely important role for a G-residue in position 5 is the triple exchange of pHL1126 which retains low promoter activity inspite of an A residue in that position. Due to altogether six concerted exchanges in positions 3, 5 and 8 as well as $\bar{3}$, $\bar{5}$ and $\bar{8}$ from the 5' and 3' end of the vRNA sequence the pHL1126 vRNA panhandle structure is indeed nearly equivalent to a wild-type cRNA panhandle, with the exception of an unpaired adenosine being present in position 10 of pHL1126 while an unpaired uridine at position $\bar{10}$ is part of the wild-type cRNA structure. This correlation may indicate a correct structure in pHL1126 for several other residues of (minor) importance which, therefore, apparently allows to compensate for the missing G residue in position 5, even if at a clearly reduced level of activity. - In the parallel pHL1102 series the corresponding triple exchange clone pHL1125 does not show any promoter activity; however, because of its deviation at position 5 it does not completely resemble the cRNA panhandle structure.

Mutational analysis of the panhandle bulge structure around nucleotide 10

An extra, unpaired residue in position 10 at the 5' end is a specific feature of the influenza viral RNA panhandle structure. It is causing or at least enforcing a major bulge of the structure, together with unpaired residues at position 9, and might be part of a specific recognition element of that structure by viral RNA polymerase. In order to investigate the importance of that particular structural feature, a further series of plasmid constructs has been initiated, again based on pHL1104 and its 3' terminal sequence as a reference. A perfectly matched RNA double-strand without any bulge has been achieved either by inserting an additional U residue in the 3' end sequence opposite A10 (pHL1140) or by deleting the A10 residue from the 5' sequence (pHL1152). Finally, a bulge of opposite direction was created in the panhandle structure of pHL1164 with an extra U residue in position 10 of the 3' end, and position 10 deleted from the 5' end sequence. While the latter two constructs proved inactive in the CAT assay, pHL1140 did show some promoter activity, albeit at a reduced level. We conclude from this result that a bulge in this region may not be recognized directly by viral RNA polymerase but may serve as a flexible joint between two more rigid structural elements that are involved in immediate contact with viral polymerase. The necessary RNA bending may also, but less efficiently be achieved in an A-U-basepaired structure like pHL1140, while the other two structures would not permit such type of interaction with RNA polymerase. This interpretation has also been substantiated in a further series of variations in this region.

Serial passaging of influenza virus carrying promoter mutants

All previous experiments consisted of a first measurement of viral mRNA synthesis in DNA-transfected and infected B82 cells, followed by a second measurement of viral mRNA synthesis in infected MDCK cells, after passaging of progeny virus containing supernatants. CAT expression in infected cells upon viral passaging requires packaging of pseudo-viral vRNAs, in addition to new rounds of viral mRNA synthesis in those cells leading to CAT expression again. All viral promoter mutants analysed and found active in transfected and helper-infected B82 cells also resulted in CAT expression after transfer, and consistently in equivalent ratios of activity. Packaging, therefore, cannot be correlated with any specific element in the vRNA promoter structure so far, and does not appear to be a limiting factor in constructing influenza virus mutants in this system. While CAT expression after passaging in general appeared to be increased over the levels

before passaging this might have been simply the result of different cells being used for the first and second step of CAT analysis, with MDCK being superior to B82 cells in influenza mRNA synthesis and also in progeny yields. Therefore, several experiments of serial passage have been performed using pHL1104 derived influenza supernatants and others, in MDCK cells. In these serial passages, always done using aliquots of supernatants harvested eight hours after infection for further transfer, a stepwise increase of CAT expression is observed (Fig. 2). Apparently the superior performance of viral RNA promoters carrying sequence deviations according to pHL1104 is not only true for viral mRNA synthesis, but also for viral RNA replication.

Therefore, mutant viral RNAs of this character become accumulated and effectively selected in further passaging, while packaging may be a neutral event in this regard, at least for the variants analysed here.

Serial passaging extended

During further passaging of supernatants the CAT containing influenza segment carrying the mutationally altered viral promotor sequences became accumulated in a stepwise manner in the population of progeny viruses. In order to demonstrate this effect on the level of individual viruses being transferred we isolated in three independent experiments 50 to 85 plaques each after a third round of passage on MDCK cells. Each cell lysate obtained for the individual plaques was assayed for CAT activity according to the standard protocol. While in two of the experiments the fraction of CAT positive plaques was in the range of 4 to 8% (1 out of 50, 4 out of 40 plaques) in one of these series this fraction amounted to 47% (19 of 40 plaques). Both of these results demonstrate a substantial increase over the initial fraction of CAT-segment containing virus, which may be calculated to be in the range of 10^{-5} or at most 10^{-4} , and slight variations in the conditions of growth during three steps of transfer may precipitate to result in the observed differences of CAT positive plaques. While every CAT positive plaque demonstrates the amplification of nine (not eight) viral RNA segments present in the initially infected cell, this may have resulted from a single virus carrying nine or more RNA segments or from coinfection by two defective viruses able to complement each other.

Necessarily, accumulation of a pseudo-viral segment not contributing to viral growth will, in further steps, become lethal to viral growth, even if a majority of virions may contain an average of eleven rather than eight RNA segments (Hsu et. al. 1987). Packaging of viral RNA-segments based on a general packaging signal identical for all eight segments and realized via a specific interaction chain: vRNA panhandle structure - viral RNA polymerase - viral NP protein - viral M1 protein will reflect the pools of the various vRNA segments in infected cells, and therefore may be biased towards an RNA segment superior in replication and overrepresented in that pool. Biased replication and packaging will, however, lead to accumulation of lethal viral particles due to an imbalance between the eight (or nine) viral RNA segments. This prediction is borne out in continuing the viral passage of pHL1104 derived influenza supernatants beyond step three as exemplified in Fig. 2. While CAT expression based on transcription of the pHL1104 derived pseudoviral RNA segments is increased further up to the fifth passage the number of viable viruses reaches a maximum already after the second step of viral passaging, thus demonstrating the continuous accumulation of an over-replicated foreign segment, based on a superior panhandle sequence.

At a stage representing the third or fourth passage as displayed in Fig. 2 a virus preparation obtained in this way can be regarded as the equivalent of an attenuated viral strain. While the concentration of attenuated virus particles that can be achieved in this way may appear to be limited a stage equivalent to passage 4 in Fig. 2 may be delayed upon coinfection with wildtype helper virus during first or second steps of transfer, and considerably increased concentrations of attenuated virus preparations might be achievable in this way.

pHL1104-mediated high-rate expression of foreign proteins can also be used (after two or more steps of amplification via serial passaging on MDCK cells) for high rate synthesis of foreign proteins in embryonated chicken eggs, following a general method of preparation of viral stocks as used for influenza and other viruses, i. e. injection of virus suspensions into the yolk sac. Protein preparations isolated from those embryonated and infected cells will be glykosylated and modified in other ways according to their origination from eukaryotic cells.

A second method of influenza virus attenuation has been achieved via cleavage of either one of the influenza viral RNAs, preferrably the M or NP gene (segments 7

or 5), via ribozyme hydrolysis in a specialised mode of action. The ribozyme RNAs which may be covalently inserted into the pSV2neo early mRNA, located between the neomycin resistance gene and the small t intron sequence originating from SV40 viral DNA, or expressed from similar expression cassettes, are directed
5 against the 5' end sequence of segment 7 (or another of the influenza vRNA segments).

During initiation of mRNA synthesis the 5' terminal sequence which is involved in formation of the panhandle structure is at first covered by viral RNA polymerase in association with that double stranded promoter region. It will, however, become
10 single-stranded and free of protein, since the polymerase molecule will start transcription at the 3' end and move along the 3' template sequence while synthesising a perfectly hybridizing 5' RNA daughter strand, superior in that regard to the parental panhandle 5' segment. Ribozyme RNAs which may be inhibited in their activity either by RNA substrates involved in double strand
15 formation or if RNA substrates are covered by protein, have been directed with a 3' complementary sequence towards that protein-free 5' sequence of the substrate vRNA molecule for initiation of hybridization, which then will be extended across the entire complementary region of approximately 100 nucleotides i. e. well into the vRNA sequence initially covered by NP protein.

20 A second feature of the ribozyme RNAs as applied for inactivation of influenza vRNA molecules is their double-unit hammerhead character, directed against not one, but two close GUY cleavage sites, e. g. GUU₁₆ and GUU₃₆ in segment 7 or GUC₃₀ and GUC₄₈ in segment 5, which are also known to be invariable in sequence comparisons of influenza isolates.

25 Both features of anti-influenza ribozymes as pointed out contribute to a reduction of typically two logs (up to three logs) in production of viral progeny in template ribozyme DNA transfected cells as compared to infection of mock transfected cells, both at moi 1 and 20 h after Lipofectamin-DNA treatment. Ribozyme treatment can be applied after two or three rounds in MDCK cells of pHL1104-
30 promoted amplification of a pseudo-viral RNA segment originating from RNA-Polymerase I transcription, in the presence of helper virus as used in initial superinfection.

In a simple version ribozyme treatment as described above is employed as a selection technique. Here, its application is appropriate if the pseudoviral RNA is indeed a (foreign) influenza segment carrying particular mutations but capable in principle to act as a functional substitute for the helper viral segment destroyed by ribozyme cleavage. For that purpose the substitute viral segment to be selected in that procedure has to be mutagenized in advance at the two cleavage sites indicated above in order to become resistant against ribozyme interference. In another application ribozyme cleavage of helpervirus vRNA can be used for attenuation of recombinant influenza virus preparations. Here, the pseudo-viral RNA segment may be designed in a way which renders it incapable to substitute for a helpervirus gene. Therefore, viral passaging into ribozyme template DNA transfected cells would lead to an abortive infection only, because of ribozyme mediated destruction of an important viral gene, if its gene product would not be added for complementation via expression from a cDNA construct which is also DNA-transfected into the cell together with ribozyme-expressing DNA 20 h before viral infection. In this way viral progeny is obtained that is attenuated because of ribozyme cleavage of one of the vRNA segments, and effectively that segment is missing in the virions because it can no longer be packaged. Viral preparations obtained in this way are capable of only one round of infection because of their inherent M1 + M2 protein complementation, and therefore are suited for vaccination purposes. Animal infection with progeny virus as isolated after the ribozyme attenuation step results in abortive infection, but viral proteins synthesized in infected cells are able to induce B-cell and T-cell responses in such animals.

In influenza viral RNA synthesis parental negative-strand vRNA is copied into plus-strand cRNA, which again is copied into progeny vRNA, from the first to the last nucleotide. This amplification of viral RNAs, however, proceeds in an inherently asymmetric way, since vRNA molecules are synthesized in excess over cRNA molecules. This result is consistent with the idea that cRNA carries a promoter structure more active in binding viral RNA polymerase and in initiation of RNA synthesis, i.e. 'stronger' than does vRNA. While at first simply the two 3' ends of single-stranded vRNA and cRNA templates have been implicated as promoter sequences, the detection of double-stranded panhandle structures involving both ends of the vRNA sequence in virions (Hsu et al., 1987) suggested more complicated substrates for RNA polymerase binding and initiation of

daughter-strand synthesis. A slightly different panhandle structure has also been observed with model vRNA molecules in the absence of viral proteins *in vitro* (Baudin et al., 1994), possibly calling for a structural change upon viral RNA polymerase binding, i.e. a bulge may be shifted from position 4 to position 10 in that reaction (see Fig. 1). While originally several of the RNA polymerase / vRNA binding experiments *in vitro* appeared to show recognition only of 3' end oligonucleotides, this has since been shown to be an artifact after pure, recombinant viral polymerase free of residual RNA became available, instead of enzyme preparations from virions. Under these conditions RNA polymerase binding to viral RNA as well as endonucleolytic cleavage of cellular mRNAs by subunit PB2 was observed to depend on vRNA 5' *plus* 3' terminal sequence binding, with even higher affinity for the 5' non-template segment (Hagen et al., 1994; Tiley et al., 1994).

Different from the employment of both vRNA and cRNA promoter structures in replication physiologically only vRNA promoters will also serve in initiation of viral mRNA synthesis according to the cap-snatching mechanism (Plotch et al., 1979; Braam et al., 1983). While it has been claimed that cRNA promoters would not have the capacity to act according to this scheme (Tiley et al., 1994), the failure to observe viral antisense mRNA molecules may simply reflect the inavailability of cRNA molecules early in infection, i.e. in the absence of surplus viral NP protein, and small amounts of such molecules might even have gone undetected. In this invention we describe a mutagenizational analysis of the vRNA promoter structure *in vivo* which in approaching the structure of the cRNA promoter via three nucleotide exchanges shows considerably improved activity in viral mRNA synthesis over vRNA promoter wild-type levels. Continuing increase of viral CAT mRNA expression during consecutive steps of viral passaging suggests that the same vRNA promoter mutants also show increased activity in cRNA synthesis, both in accordance with the idea that the cRNA promoter structure might be 'stronger' than the vRNA promoter, also in initiation of viral mRNA synthesis.

Additional variations of the 5' terminal sequence clearly indicate the major importance of a G residue in position 5, irrespective of complementarity or not to position 5 at the 3' end. The unique role of this G residue has been observed before in a serial mutagenizational analysis (Li and Palese, 1992). According to

both data guanosine residue 5 may be involved in single-strand binding of RNA polymerase as has indeed been observed for the non-template strand terminal segment (Tiley et al., 1994). While panhandle double-strand structures are likely to constitute the initial RNA polymerase binding substrate a partial separation of template and non-template strands is expected to take place consecutively resulting in a 'forked structure' such as proposed by Fodor et al. (1994). Specific and tight binding of RNA polymerase in this structure may predominantly be oriented towards sequence elements in the non-template strand, since the growing point of RNA synthesis will have to move along the entire template strand following its initiation. It is, therefore, possible that such a binding interaction survives most or all of an individual round of mRNA synthesis as has been proposed (Tiley et al., 1994).

The triple nucleotide exchanges as introduced in vRNA molecules derived from pHL1104 templates will create three additional basepairs able to stabilize the panhandle structure in general, but more specifically they will favor a bulged adenosine 10 over the bulged adenosine 4 conformation as observed for the wild-type sequence *in vitro* (Baudin et al., 1994). Since the changes introduced here lead to a considerable enhancement of promoter activity we propose that a bulged 10 conformation may be the structure underlying the vRNA / polymerase binding reaction, which otherwise would have to be achieved only as a result of that interaction. A bulged 10 adenosine residue may constitute a kind of flexible joint or angular kink which in turn suggests two major, structurally stable binding sites to the left and right of this element. One of these sites has to be the double-stranded sequence element of (in average) six basepairs extending from positions 11 to 16 and $\overline{10}$ to $\overline{15}$, respectively. While the distal three basepairs are known to be variable for the various RNA segments, basepair 13/ $\overline{12}$ has been shown to be exchangeable experimentally, and also the number of basepairs has been reduced to four without complete loss of function (Luo et al., 1991). With all of these data it seems clear that the main recognition element in this region is an RNA double-strand of certain stability, while it remains possible that residue 12 guanosine and potentially others are also recognized individually within that structure. A major second binding element for RNA polymerase on the other side relative to position 10 is less evident, but may be located in a distance of nearly one helical turn in the de-bulged region around position 4, since direct contacts are suggested by that initial conformational interaction, and also by the specific requirement of a

guanosine residue in position 5, which is likely to interact not only during, but also before partial strand separation in that region, i.e. in the panhandle as well as the forked structure. While an extra adenosine residue in position 10 may be optimal for creating a correctly shaped bulge in this region of RNA, structural variants are possible in this regard (see pHL 1140) which excludes direct interactions between RNA polymerase and residues constituting that bulge.

In summary we are proposing a model (see Fig. 1) of consecutive steps of interaction between a vRNA or cRNA promoter structure and viral RNA polymerase:

10 bulged 4 panhandle → bulged 10 panhandle / polymerase → forked RNA / polymerase (bound to 5G and ds element 11-16) → initiation of RNA synthesis (recognition of 3' end of template).

Attenuation of influenza viruses for preparation of a live nasal vaccine relies on two mechanisms: 1) preferential amplification of a recombinant viral segment carrying the pHL1104 promoter mutation, which will increase its rate in packaged viral RNP particles and indirectly decrease that of the eight helper virus RNP particles. This competition results in an increase of defective viruses from which one or more of the regular viral gene segments are missing. 2) Sequence specific ribozyme cleavage of one or more helper virus RNA segments, if compensated through gene product expression for functional complementation. This dual interaction will result in virus progeny, which is capable of only one round of infection, abortive because of the missing viral protein(s) that are required for their propagation. - Ribozyme cleavage of one out of two sister viral gene segments, sensitive and (artificially) insensitive for its hydrolysis may also be used (repeatedly) for selection purposes, including selection for viral gene constructs expressed via RNA polymerase I transcription.

MATERIALS AND METHODS

Plasmid constructions

30 Plasmids with mutated vRNA and/or mutated cRNA promoter sequences are derivatives of pHL926 (Zobel et al., 1993; Neumann et al., 1994). In pHL926 a hybrid CAT cDNA with flanking non-coding sequences derived from influenza vRNA segments has been precisely inserted in antisense orientation between

mouse rDNA promoter and terminator sequences. The CAT reporter gene in this way has been introduced by exactly replacing the coding sequence for hemagglutinin, retaining the untranslated viral 5' and 3' sequences of segment 4.

5 vRNA 5' end mutations were created by PCR, using a general primer hybridizing to a position in the flanking rDNA promoter sequence, and a specific primer carrying the desired nucleotide substitution to be introduced in the viral terminal sequence. The polymerase chain reaction products were first digested by the restriction enzymes *Bgl*II and *Spe*I, inserted into the left boundary position by exchanging the segment between these appropriate restriction sites in pHL926, and
10 finally confirmed in their constitution by DNA Sanger sequencing.

Generation of vRNA 3' end mutations followed the same general scheme at the right boundary. PCR products were obtained by using a general primer complementary to a CAT gene internal sequence position, and a specific primer with appropriate nucleotide exchanges inserted into its sequence. Following
15 digestion with restriction enzymes *Nco*I and *Sca*I, the PCR products were cloned into *Nco*I- and *Sca*I(partially)-digested plasmid pHL926. Any PCR derived sequences were investigated by DNA sequencing.

For constructs with both 5' end and 3' end mutations in combination, 5' variation containing fragments were obtained by *Bgl*II and *Spe*I restriction and inserted into
20 the appropriate 3' terminal variation plasmids.

Cells and viruses

Influenza A/FPV/Bratislava viruses were grown in NIH3T3 cells. For transfection and passaging experiments B82 cells (a mouse L cell line) and MDCK cells were used.

25 Lipofectamin DNA transfection and influenza virus helper infection

For DNA transfection 10^7 B82 cells were used. 5 µg of plasmid DNA were mixed with 60 µg of Lipofectamin (LipofectaminTM, GIBCO/BRL) in serum-free medium and incubated at room temperature for 10-15 min. This mixture was added to the cells washed twice with serum-free medium, and the incubation with
30 Lipofectamin/DNA was continued for 1 hr. After further incubation with DMEM medium for 1 hr the transfected B82 cells were infected with influenza

A/FPV/Bratislava at a multiplicity of infection of 0.01 to 1 for another 30-60 min. Further incubation was performed with DMEM medium.

Passaging of virus containing supernatants

5 Under standard conditions 8 hr after influenza infection (at moi 0.1 to 1) cells were harvested for CAT assays, and supernatants were collected and spun down at 1200 rpm for 5 min for removal of cell debris.

10 Aliquots of virus containing cleared supernatants were used for plaque tests, and another aliquot was adsorbed to 10^7 MDCK cells for 30-60 min for further passaging. Again 8 hr after infection the CPE was verified, and cells and supernatants were collected and treated as before.

CAT assay

15 Cell extracts were prepared as described by Gorman et al. (1982). CAT assays were done with [14 C]chloramphenicol or fluorescent-labeled chloramphenicol (borondipyrromethane difluoride fluorophore; FLASH CAT Kit, Stratagene) as substrates.

20 For [14 C]chloramphenicol the assay mixture contained: 0.1 μ Ci [14 C]chloramphenicol, 20 μ l 4 mM Acetyl-CoA, 25 μ l 1 M Tris-HCl (pH 7.5) and 50 μ l of cell lysate in a total volume of 150 μ l. The assay mixture for the fluorescent-labeled substrate contained (in a final volume of 80 μ l): 10 μ l 0.25 M Tris-HCl (pH 7.5), 10 μ l 4 M Acetyl-CoA, 10 μ l fluorescent-labeled chloramphenicol, and 50 μ l of cell lysate. After an incubation time of 16 hr the reaction products were separated by chromatography and either autoradiographed or visualized by UV illumination and photography.

REFERENCES

- 25 Baudin,F., Bach,C., Cusack,S. and Ruigrok,R.W.H. (1994) *The EMBO J.*, **13**, 3158-3165.
Braam,J., Uhmanen,I. and Krug,R. (1983) *Cell*, **34**, 609-618.
Compans,R.W., Content,J. and Duesberg,P.H. (1972) *J. Virol.*, **10**, 795-800.
Fodor,E., Seong,B.L. and Brownlee,G.G. (1993) *J. Gen. Virol.*, **74**, 1327-1333.
30 Fodor,E., Pritlove,D.C. and Brownlee,G.G. (1994) *J. Virol.*, **68**, 4092-4096.
Gorman,M., Moffat,L. and Howard,B. (1982) *Mol. Cell Biol.*, **2**, 1044-1057.

- Hagen,M., Chung,T.D.Y., Butcher,J.A. and Krystal,M. (1994) *J. Virol.*, **68**, 1509-1515.
- Honda,A., Ueda,K., Nagata,K. and Ishihama,A. (1987) *J. Biochem.*, **102**, 1241-1249.
- 5 Hsu,M., Parvin,J.D., Gupta,S., Krystal,M. and Palese,P. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8140-8144.
- Li,X. and Palese,P. (1992) *J. Virol.*, **66**, 4331-4338.
- Luo,G., Luytjes,W., Enami,M. and Palese,P. (1991) *J. Virol.*, **65**, 2861-2867.
- Martin,J., Albo,C., Ortin,J., Melero,J.A. and Portela,A. (1992) *J. Gen. Virol.*, **73**,
10 1855-1859.
- Neumann,G., Zobel,A. and Hobom,G. (1994) *Virology*, **202**, 477-479.
- Plotch,S., Bouloy,M. and Krug,R.M. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 1618-1622.
- Seong,B.L. and Brownlee,G.G. (1992) *J. Virol.*, **73**, 3115-3124.
- 15 Shapiro,G. and Krug,R. (1988) *J. Virol.*, **62**, 2285-2290.
- Tiley,L.S., Hagen,M., Matthews,J.T. and Krystal,M. (1994) *J. Virol.*, **68**, 5108-5116.
- Yamanaka,K., Ogasawara,N., Yoshikawa,H., Ishihama,A. and Nagata,K. (1991) *Proc. Natl. Acad. Sci. USA*, **88**: 5369-5373.
- 20 Zobel,A., Neumann,G. and Hobom,G. (1993) *Nucl. Acids Res.*, **21**, 3607-3614.

LEGENDS TO FIGURES

Fig. 1. Proposed scheme of consecutive conformational steps occurring prior to initiation of viral mRNA synthesis in influenza vRNA, in wildtype and pHL1104 derived mutant sequences. Positions of triple mutation in pHL1104 vRNA are indicated in bold and larger size letters.

25

(A) Free RNA panhandle structure, bulged at position 4 (wildtype vRNA; Baudin et al, 1994) or at position 10 (mutant vRNA). (B) Bulged 10 panhandle structures after binding of viral RNA polymerase; proposed protein binding positions marked by underlinements. (C) Forked structures of partial strand separation. (D) Initiation of viral mRNA synthesis via hybridization of capped primer oligonucleotide.

30

Fig. 2. Serial passaging of pHL 1104-derived progeny viruses.

10⁷ B82 cells were transfected with 5μ of pHL 1104 DNA (in 60 μg Lipofectamin) and infected two hours later with influenza A/FPV/Bratislava (m.o.i:1). 8 hr post-infection the cells were assayed for plaque forming units (dark column) and for CAT activity (hatched column) After sedimentation an other
5 aliquot of the supernatant was adsorbed to 10⁷ MDCK cells. Further rounds of passaging were done equivalently by harvesting the cells 8 hr after infection for assaying CAT activities, whereas an aliquot of the supernatant was always adsorbed to fresh MDCK cells. Numbers of serial passages are indicated at the bottom. Plaque forming units per ml refer to the left ordinate, CAT expression
10 rates (relative to primary imfection levels) to the ordinate on the right.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: Bayer AG
(B) STREET: Bayerwerk
(C) CITY: Leverkusen
(E) COUNTRY: Deutschland
(F) POSTAL CODE (ZIP): 51368
(G) TELEPHONE: (0)214-3061455
(H) TELEFAX: (0)214-303482

(ii) TITLE OF INVENTION: Vaccination virus, method of making it and
pharmaceutical composition comprising that virus

(iii) NUMBER OF SEQUENCES: 8

(iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30B (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5241 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Influenza virus, RNA sequence
(C) INDIVIDUAL ISOLATE: pHL926

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

CTGGCGGCAG TAGCGCGGTG GTCCACCTG ACCCATGCC GAACTCAGAA GTGAAACGCC	60
GTAGCGCCGA TGGTAGTGTG GGGTCTCCCC ATGCGAGAGT AGGGAAGTGC CAGGCATCAA	120
ATAAAACGAA AGGCTCAGTC GAAAGACTGG GCCTTTCGTT TTATCTGTTG TTTGTCGGTG	180
AACGCTCTCC TGAGTAGGAC AAATCCGCCG GGAGCGGATT TGAACGTTGC GAAGCAACGG	240

CCCCGAGGGT GCGGGGCAGG ACGCCCGCCA TAACTGCCA GGCATCAAAT TAAGCAGAAG 300
GCCATCCTGA CGGATGGCCT TTTTGC GTTT CTACAAACTC TTTTGTTTAT TTTTCTAAAT 360
ACATTCAAAT ATGTATCCGC TCATGAGACA ATAACCCTGA TAAATGCTTC AATAATATTG 420
AAAAAGGAAG AGTATGAGTA TTCAACATTT CCGTGTGCGC CTTATTCCTT TTTTTCGGGC 480
5 ATTTTGCCTT CCTGTTTTTG CTCACCCAGA AACGCTGGTG AAAGTAAAAG ATGCTGAAGA 540
TCAGTTGGGT GCACGAGTGG GTTACATCGA ACTGGATCTC AACAGCGGTA AGATCCTTGA 600
GAGTTTTTCGC CCCGAAGAAC GTTTTCCAAT GATGAGCACT TTTAAAGTTC TGCTATGTGG 660
CGCGGTATTA TCCCGTGTTG ACGCCGGGCA AGAGCAACTC GGTGCGCGCA TACACTATTC 720
TCAGAATGAC TTGGTTGAGT ACTCACCAGT CACAGAAAAG CATCTTACGG ATGGCATGAC 780
10 AGTAAGAGAA TTATGCAGTG CTGCCATAAC CATGAGTGAT AACACTGCGG CCAACTTACT 840
TCTGACAACG ATCGGAGGAC CGAAGGAGCT AACCGCTTTT TTGCACAACA TGGGGGATCA 900
TGTAAGTCGC CTTGATCGTT GGGAAACCGGA GCTGAATGAA GCCATACCAA ACGACGAGCG 960
TGACACCACG ATGCCTGCAG CAATGGCAAC AACGTTGCGC AACTATTAA CTGGCGAACT 1020
ACTTACTCTA GCTTCCCGGC AACAATTAAT AGACTGGATG GAGGCGGATA AAGTTGCAGG 1080
15 ACCACTTCTG CGCTCGGCCC TTCCGGCTGG CTGGTTTATT GCTGATAAAT CTGGAGCCGG 1140
TGAGCGTGGG TCTCGCGGTA TCATTGCAGC ACTGGGGCCA GATGGTAAGC CCTCCCGTAT 1200
CGTAGTTATC TACACGACGG GGAGTCAGGC AACTATGGAT GAACGAAATA GACAGATCGC 1260
TGAGATAGGT GCCTCACTGA TTAAGCATTG GTAAGTGTCA GACCAAGTTT ACTCATATAT 1320
ACTTTAGATT GATTTAAAAC TTCATTTTTA ATTTAAAAGG ATCTAGGTGA AGATCCTTTT 1380
20 TGATAATCTC ATGACCAAAA TCCCTTAACG TGAGTTTTTCG TTCCACTGAG CGTCAGACCC 1440
CGTAGAAAAG ATCAAAGGAT CTTCTTGAGA TCCTTTTTTT CTGCGCGTAA TCTGCTGCTT 1500
GCAAACAAAA AAACCACCGC TACCAGCGGT GGTGTTGTTT CCGGATCAAG AGCTACCAAC 1560
TCTTTTCCG AAGGTAAGT GCTTCAGCAG AGCGCAGATA CCAAATACTG TCCTTCTAGT 1620
GTAGCCGTAG TTAGGCCACC ACTTCAAGAA CTCTGTAGCA CCGCCTACAT ACCTCGCTCT 1680
25 GCTAATCCTG TTACAGTGG CTGCTGCCAG TGGCGATAAG TCGTGTCTTA CCGGGTTGGA 1740
CTCAAGACGA TAGTTACCGG ATAAGGCGCA GCGGTCGGG TGAACGGGG GTTCGTGCAC 1800
ACAGCCCAGC TTGGAGCGAA CGACCTACAC CGAACTGAGA TACCTACAGC GTGAGCATTG 1860

	AGAAAGCGCC	ACGCTTCCCG	AAGGGAGAAA	GGCGGACAGG	TATCCGGTAA	GCGGCAGGGT	1920
	CGGAACAGGA	GAGCGCACGA	GGGAGCTTCC	AGGGGGAAAC	GCCTGGTATC	TTTATAGTCC	1980
	TGTCGGGTTT	CGCCACCTCT	GACTTGAGCG	TCGATTTTTG	TGATGCTCGT	CAGGGGGGCG	2040
	GAGCCTATGG	AAAAACGCCA	GCAACGCGGC	CCGAGATGCG	CCGCGTGCGG	CTGCTGGAGA	2100
5	TGGCGGACGC	GATGGATATG	TTCTGCCAAG	GGTTGGTTTG	CGCATTCACA	GTTCTCCGCA	2160
	AGAATTGATT	GGCTCCAATT	CTTGGAGTGG	TGAATCCGTT	AGCGAGGTGC	CGCCGGCTTC	2220
	CATTCAGGTC	GAGGTGGCCC	GGCTCCATGC	ACCGCGACGC	AACGCGGGGA	GGCAGACAAG	2280
	GTATAGGGCG	GCGCCTACAA	TCCATGCCAA	CCCGTTCCAT	GTGCTCGCCG	AGGCGGCATA	2340
	AATCGCCGTG	ACGATCAGCG	GTCCAGTGAT	CGAAGTTAGG	CTGGTAAGAG	CCGCGAGCGA	2400
10	TCCTTGAAGC	TGTCCCTGAT	GGTCGTCATC	TACCTGCCTG	GACAGCATGG	CCTGCAACGC	2460
	GGGCATCCCG	ATGCCGCCGG	AAGCGAGAAG	AATCATAATG	GGGAAGGCCA	TCCAGCCTCG	2520
	CGTCGCGAAC	GCCAGCAAGA	CGTAGCCCAG	CGCGTCGGCC	AGCTTGCAAT	TCGCGCTAAC	2580
	TTACATTAAT	TGCGTTGCGC	TCACTGCCCC	CTTTCCAGTC	GGGAAACCTG	TCGTGCCAGC	2640
	TGCATTAATG	AATCGGCCAA	CGCGCGGGGA	GAGGCGGTTT	GCGTATTGGG	CGCCAGGGTG	2700
15	GTTTTTCTTT	TCACCACTGA	GACGGGCAAC	AGCTGATTGC	CCTTCACCGC	CTGGCCCTGA	2760
	GAGAGTTGCA	GCAAGCGGTC	CACGCTGGTT	TGCCCCAGCA	GGCGAAAATC	CTGTTTGATG	2820
	GTGGTTGACG	GCGGGATATA	ACATGAGCTG	TCTTCGGTAT	CGTCGTATCC	CACTACCGAG	2880
	ATATCCGCAC	CAACGCGCAG	CCCGGACTCG	GTAATGGCGC	GCATTGCGCC	CAGCGCCATC	2940
	TGATCGTTGG	CAACCAGCAT	CGCAGTGGGA	ACGATGCCCT	CATTCAGCAT	TTGCATGGTT	3000
20	TGTTGAAAAC	CGGACATGGC	ACTCCAGTCG	CCTTCCCGTT	CCGCTATCGG	CTGAATTTGA	3060
	TTGCGAGTGA	GATATTTATG	CCAGCCAGCC	AGACGCAGAC	GCGCCGAGAC	AGAACTTAAT	3120
	GGGCCCCCGG	TGACAGGGAC	AGAGAGGGCT	TCTGGAGGAA	AAAAGAAAAA	AAAAAAAAAA	3180
	GATCCAAAGC	TCCAGGGCGA	GCTCGAATTC	CCCGGTAAAG	CCGCTTAAGA	CATTCCCCT	3240
	CTTACACATC	CCAGCCCTGA	AAAAGGGCAT	CAAAATAAAC	CACACCTATG	GTGTATGCAT	3300
25	TTACGTTGAC	ACCATCGAAT	GGTGCAAAAC	CTTTCGCGGT	ATGGCATGAT	AGCGCCCGGA	3360
	AGAGAGTCAA	TTCAGGGTGG	TGAATGTGAA	ACCAGTAACG	TTATACGATG	TCGCAGAGTA	3420
	TGCCGGTGTC	TCTTATCAGA	CCGTTTCCCG	CGTGGTGAAC	CAGGCCAGCC	ACGTTTCTGC	3480

GAAAACGCGG GAAAAAGTGG AAGCGGCGAT GCGGAGCTG AATTACATTC CCAACCGCGT 3540
 GGCACAACAA CTGGCGGGCA AACAGTCGTT GCTGATTGGC GTTGCCACCT CCAGTCTGGC 3600
 CCTGCACGCG CCGTCGCAA TTGTCGCGC GATTAAATCT CGCGCCGATC AACTGGGTGC 3660
 CAGCGTGGTG GTGTCGATGG TAGAACGAAG CGGCGTCGAA GCCTGTAAAG CGGCGGTGCA 3720
 5 CAATCTTCTC GCGCAACGCG TTCAGTGGGC TAGATCTACC GCGGTAGATC ATTAACATC 3780
 CGCTGGATGA CCAGGATGCC ATTGCTGTGG AAGCTGCCTG CACTAATGTT CCGGCGTTAT 3840
 TTCTTGATGT CTCTGACCAG ACACCCATCA ACAGTATTAT TTTCTCCCAT GAAGACGGTA 3900
 CGCGACTGGG CGTGGAGCAT CTGGTCGCAT TGGGTCACCA GCAAATCGCG CTGTTAGCGG 3960
 GCCCTGTACG TCTGAGGCCG AGGGAAAGCT ATGGGCGCGG TTTTCTTTCA TTGACCTGTC 4020
 10 GGTCTTATCA GTTCTCCGGG TTGTCAGGTC GACCAGTTGT TCCTTTGAGG TCCGGTTCTT 4080
 TTCGTTATGG GGTCAATTTT GGGCCACCTC CCCAGGTATG ACTTCCAGGT ATTCTCTGTG 4140
 GCCTGTCACT TTCCTCCCTG TCTCTTTTAT GCTTGTGATC TTTTAGATCT GGTCTATTG 4200
 GACCTGGAGA TAGGTAGTAG AAACAAGGGT GTTTTTAAAT ACTAGTACAT TACGCCCCGC 4260
 CCTGCCACTC ATCGCAGTAC TGTTGTAATT CATTAAAGCAT TCTGCCGACA TGGAAGCCAT 4320
 15 CACAGACGGC ATGATGAACC TGAATCGCCA GCGGCATCAG CACCTTGTCG CCTTGCGTAT 4380
 AATATTTGCC CATGGTGAAA ACGGGGGCGA AGAAGTTGTC CATATTGGCC ACGTTTAAAT 4440
 CAAAACTGGT GAAACTCACC CAGGGATTGG CTGAGACGAA AAACATATTC TCAATAAACC 4500
 CTTTAGGGAA ATAGGCCAGG TTTTCACCGT AACACGCCAC ATCTTGCGAA TATATGTGTA 4560
 GAAACTGCCG GAAATCGTCG TGGTATTAC TCCAGAGCGA TGAAAACGTT TCAGTTTGCT 4620
 20 CATGGAAAAC GGTGTAACAA GGGTGAACAC TATCCCATAT CACCAGCTCA CCGTCTTTCA 4680
 TTGCCATACG GAATCCGGA TGAGCATTCA TCAGGCGGGC AAGAATGTGA ATAAAGGCCG 4740
 GATAAACTT GTGCTTATTT TTCTTTACGG TCTTTAAAAA GGCCGTAATA TCCAGCTGAA 4800
 CGGTCTGGTT ATAGGTACAT TGAGCAACTG ACTGAAATGC CTCAAAATGT TCTTTACGAT 4860
 GCCATTGGGA TATATCAACG GTGGTATATC CAGTGATTTT TTTCTCCATG ATTAATAGAA 4920
 25 TTATCCCCTG TTTCTACTCC CCCCCAACTT CGGAGGTGCA CCAGTACTCC GGGCGAAACT 4980
 TTGTTTTTTT TTTTCCCCC GATGCTGGAG GTCGACCAGA TGTCCGAAAG TGTCCCCCCC 5040
 CCCCCCCCCC CCGGCGCGG AACGGCGGGG CCACTCTGGA CTCTTTTTTT TTTTTTTTTT 5100

TTTTTTTTTG GGGATCCTCT AGAGTCGACC TGCAGCCCAA GCTAGCGGCC GCTAGCTTCT 5160
GTTTGGCGG ATGAGAGAAG ATTTTCAGCC TGATACAGAT TAAATCAGAA CGCAGAAGCG 5220
GTCTGATAAA ACAGAATTTG C 5241

(2) INFORMATION FOR SEQ ID NO: 2:

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5241 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular
- 10 (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
15 (A) ORGANISM: Influenza virus, RNA sequence
(C) INDIVIDUAL ISOLATE: pHL1104

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

CTGGCGGCAG TAGCGCGGTG GTCCACCTG ACCCATGCC GAACTCAGAA GTGAAACGCC 60
GTAGCGCCGA TGGTAGTGTG GGGTCTCCCC ATGCGAGAGT AGGGAAGTGC CAGGCATCAA 120
ATAAAACGAA AGGCTCAGTC GAAAGACTGG GCCTTTCGTT TTATCTGTTG TTTGTCGGTG 180
20 AACGCTCTCC TGAGTAGGAC AAATCCGCCG GGAGCGGATT TGAACGTTGC GAAGCAACGG 240
CCCGGAGGGT GGCGGGCAGG ACGCCCGCCA TAAACTGCCA GGCATCAAAT TAAGCAGAAG 300
GCCATCCTGA CGGATGGCCT TTTTGC GTTT CTACAACTC TTTTGTTTAT TTTTCTAAAT 360
ACATTCAAAT ATGTATCCGC TCATGAGACA ATAACCCTGA TAAATGCTTC AATAATATTG 420
AAAAAGGAAG AGTATGAGTA TTCAACATTT CCGTGTGCGC CTTATTCCCT TTTTGC GGCG 480
25 ATTTTGCCTT CCTGTTTTTG CTCACCCAGA AACGCTGGTG AAAGTAAAAG ATGCTGAAGA 540
TCAGTTGGGT GCACGAGTGG GTTACATCGA ACTGGATCTC AACAGCGGTA AGATCCTTGA 600
GAGTTTTTCGC CCCGAAGAAC GTTTTCCAAT GATGAGCACT TTAAAGTTC TGCTATGTGG 660
CGCGGTATTA TCCCGTGTTG ACGCCGGGCA AGAGCAACTC GGTGCGCCGCA TACACTATTC 720
TCAGAATGAC TTGGTTGAGT ACTCACCAGT CACAGAAAAG CATCTTACGG ATGGCATGAC 780

	AGTAAGAGAA TTATGCAGTG CTGCCATAAC CATGAGTGAT AACACTGCGG CCAACTTACT	840
	TCTGACAACG ATCGGAGGAC CGAAGGAGCT AACCGCTTTT TTGCACAACA TGGGGGATCA	900
	TGTAACCTCGC CTTGATCGTT GGGAAACCGGA GCTGAATGAA GCCATACCAA ACGACGAGCG	960
	TGACACCACG ATGCCTTCAG CAATGGCAAC AACGTTGCGC AAACATATTAA CTGGCGAACT	1020
5	ACTTACTCTA GCTTCCCGGC AACAATTAAT AGACTGGATG GAGGCGGATA AAGTTGCAGG	1080
	ACCACTTCTG CGCTCGGCCC TTCCGGCTGG CTGGTTTATT GCTGATAAAT CTGGAGCCGG	1140
	TGAGCGTGGG TCTCGCGGTA TCATTGCAGC ACTGGGGCCA GATGGTAAGC CCTCCCGTAT	1200
	CGTAGTTATC TACACGACGG GGAGTCAGGC AACTATGGAT GAACGAAATA GACAGATCGC	1260
	TGAGATAGGT GCCTCACTGA TTAAGCATTG GTAACGTGCA GACCAAGTTT ACTCATATAT	1320
10	ACTTTAGATT GATTTAAAAC TTCATTTTTA ATTTAAAAGG ATCTAGGTGA AGATCCTTTT	1380
	TGATAATCTC ATGACCAAAA TCCCTTAACG TGAGTTTTTCG TTCCACTGAG CGTCAGACCC	1440
	CGTAGAAAAG ATCAAAGGAT CTTCTTGAGA TCCTTTTTTT CTGCGCGTAA TCTGCTGCTT	1500
	GCAAACAAAA AAACCACCGC TACCAGCGGT GGTGTTGTTG CCGGATCAAG AGCTACCAAC	1560
	TCTTTTCCG AAGGTAAC TGCTTCAGCAG AGCGCAGATA CCAAATACTG TCCTTCTAGT	1620
15	GTAGCCGTAG TTAGGCCACC ACTTCAAGAA CTCTGTAGCA CCGCCTACAT ACCTCGCTCT	1680
	GCTAATCCTG TTACCAGTGG CTGCTGCCAG TGGCGATAAG TCGTGCTTGA CCGGGTTGGA	1740
	CTCAAGACGA TAGTTACCGG ATAAGGCGCA GCGGTGCGGC TGAACGGGGG GTTCGTGCAC	1800
	ACAGCCCAGC TTGGAGCGAA CGACCTACAC CGAACTGAGA TACCTACAGC GTGAGCATTG	1860
	AGAAAGCGCC ACGCTTCCCG AAGGGAGAAA GGCGGACAGG TATCCGGTAA GCGGCAGGGT	1920
20	CGGAACAGGA GAGCGCACGA GGGAGCTTCC AGGGGGAAAC GCCTGGTATC TTTATAGTCC	1980
	TGTCGGGTTT CGCCACCTCT GACTTGAGCG TCGATTTTTG TGATGCTCGT CAGGGGGGCG	2040
	GAGCCTATGG AAAAACGCCA GCAACGCGGC CCGAGATGCG CCGCGTGCGG CTGCTGGAGA	2100
	TGGCGGACGC GATGGATATG TTCTGCCAAG GGTGTTTGG CGCATTACCA GTTCTCCGCA	2160
	AGAATTGATT GGCTCCAATT CTTGGAGTGG TGAATCCGTT AGCGAGGTGC CGCCGGCTTC	2220
25	CATTGAGGTC GAGGTGGCCC GGCTCCATGC ACCGCGACGC AACGCGGGGA GGCAGACAAG	2280
	GTATAGGGCG GCGCCTACAA TCCATGCCAA CCCGTTCCAT GTGCTCGCCG AGGCGGCATA	2340
	AATCGCCGTG ACGATCAGCG GTCCAGTGAT CGAAGTTAGG CTGGTAAGAG CCGCGAGCGA	2400

TCCTTGAAGC TGTCCCTGAT GGTCGTCATC TACCTGCCTG GACAGCATGG CCTGCAACGC 2460
GGGCATCCCG ATGCCGCCGG AAGCGAGAAG AATCATAATG GGGAAGGCCA TCCAGCCTCG 2520
CGTCGCGAAC GCCAGCAAGA CGTAGCCCAG CGCGTCGGCC AGCTTGCAAT TCGCGCTAAC 2580
TTACATTAAT TGC GTTGC GC TCACTGCCCC CTTTCCAGTC GGGAAACCTG TCGTGCCAGC 2640
5 TGCATTAATG AATCGGCCAA CGCGCGGGGA GAGGCGGTTT GCGTATTGGG CGCCAGGGTG 2700
GTTTTTCTTT TCACCAGTGA GACGGGCAAC AGCTGATTGC CTTTACC GC CTGGCCCTGA 2760
GAGAGTTGCA GCAAGCGGTC CACGCTGGTT TGCCCCAGCA GGC GAAAATC CTGTTTGATG 2820
GTGGTTGACG GCGGGATATA ACATGAGCTG TCTTCGGTAT CGTCGTATCC CACTACCGAG 2880
ATATCCGCAC CAACGCGCAG CCCGGACTCG GTAATGGCGC GCATTGCGCC CAGCGCCATC 2940
10 TGATCGTTGG CAACCAGCAT CGCAGTGGGA ACGATGCCCT CATTCAGCAT TTGCATGGTT 3000
TGTTGAAAAC CGGACATGGC ACTCCAGTCG CTTTCCCGTT CCGCTATCGG CTGAATTTGA 3060
TTGCGAGTGA GATATTTATG CCAGCCAGCC AGACGCAGAC GCGCCGAGAC AGAACTTAAT 3120
GGGCCCCCGG TGACAGGGAC AGAGAGGGCT TCTGGAGGAA AAAAGAAAAA AAAAAAAAAA 3180
GATCCAAAGC TCCAGGGCGA GCTCGAATTC CCCGGTAAAG CCGCTTAAGA CATTCCCGCT 3240
15 CTTACACATC CCAGCCCTGA AAAAGGGCAT CAAAATAAAC CACACCTATG GTGTATGCAT 3300
TTACGTTGAC ACCATCGAAT GGTGCAAAAC CTTTCGCGGT ATGGCATGAT AGCGCCCGGA 3360
AGAGAGTCAA TTCAGGGTGG TGAATGTGAA ACCAGTAACG TTATACGATG TCGCAGAGTA 3420
TGCCGGTGTC TCTTATCAGA CCGTTTCCCG CGTGGTGAAC CAGGCCAGCC ACGTTTCTGC 3480
GAAAACGCGG GAAAAAGTGG AAGCGGCGAT GGCGGAGCTG AATTACATTC CCAACCGCGT 3540
20 GGCACAACAA CTGGCGGGCA AACAGTCGTT GCTGATTGGC GTTGCCACCT CCAGTCTGGC 3600
CCTGCACGCG CCGTCGCAA TTGTCGCGGC GATTAAATCT CGCGCCGATC AACTGGGTGC 3660
CAGCGTGGTG GTGTCGATGG TAGAACGAAG CGGCGTCGAA GCCTGTAAAG CGGCGGTGCA 3720
CAATCTTCTC GCGCAACGCG TTCAGTGGGC TAGATCTACC GCGGTAGATC ATTA ACTATC 3780
CGCTGGATGA CCAGGATGCC ATTGCTGTGG AAGCTGCCTG CACTAATGTT CCGGCGTTAT 3840
25 TTCTTGATGT CTCTGACCAG ACACCCATCA ACAGTATTAT TTTCTCCAT GAAGACGGTA 3900
CGCGACTGGG CGTGGAGCAT CTGGTCGCAT TGGGTCACCA GCAAATCGCG CTGTTAGCGG 3960
GCCCTGTACG TCTGAGGCCG AGGGAAAGCT ATGGGCGCGG TTTTCTTTCA TTGACCTGTC 4020

GGTCTTATCA GTTCTCCGGG TTGTCAGGTC GACCAGTTGT TCCTTTGAGG TCCGGTTCTT 4080
TTCGTTATGG GGTCATTTTT GGGCCACCTC CCCAGGTATG ACTTCCAGGT ATTCTCTGTG 4140
GCCTGTCACT TTCCTCCCTG TCTCTTTTAT GCTTGTGATC TTTTAGATCT GGTCCTATTG 4200
GACCTGGAGA TAGGTAGTAG AAACAAGGGT GTTTTTAAAT ACTAGTACAT TACGCCCCGC 4260
5 CCTGCCACTC ATCGCAGTAC TGTGTAAATT CATTAAGCAT TCTGCCGACA TGGAAGCCAT 4320
CACAGACGGC ATGATGAACC TGAATCGCCA GCGGCATCAG CACCTTGTCTG CCTTGCGTAT 4380
AATATTTGCC CATGGTGAAA ACGGGGGCGA AGAAGTTGTC CATATTGGCC ACGTTTAAAT 4440
CAAAACTGGT GAAACTCACC CAGGGATTGG CTGAGACGAA AAACATATTC TCAATAAACC 4500
CTTTAGGGAA ATAGGCCAGG TTTTCACCGT AACACGCCAC ATCTTGCGAA TATATGTGTA 4560
10 GAAACTGCCG GAAATCGTCG TGGTATTCAC TCCAGAGCGA TGAAAACGTT TCAGTTTGCT 4620
CATGGAAAAC GGTGTAACAA GGGTGAACAC TATCCCATAT CACCAGCTCA CCGTCTTTCA 4680
TTGCCATACG GAATTCGGGA TGAGCATTCA TCAGGCGGGC AAGAATGTGA ATAAAGGCCG 4740
GATAAACTT GTGCTTATTT TTCTTTACGG TCTTTAAAAA GGCCGTAATA TCCAGCTGAA 4800
CGGTCTGGTT ATAGGTACAT TGAGCAACTG ACTGAAATGC CTCAAAATGT TCTTTACGAT 4860
15 GCCATTGGGA TATATCAACG GTGGTATATC CAGTGATTTT TTTCTCCATG ATTAATAGAA 4920
TTATCCCCTG TTTCTACTCC CCCCCAACTT CGGAGGTCTGA CCAGTACTCC GGGCGAAACT 4980
TTGTTTTTTT TTTTCCCCC GATGCTGGAG GTCGACCAGA TGTCCGAAAG TGTCACCCCC 5040
CCCCCCCCC CCCGGCGCGG AACGGCGGGG CCACTCTGGA CTCTTTTTTT TTTTTTTTTT 5100
TTTTTTTTTG GGGATCCTCT AGAGTCGACC TGCAGCCCAA GCTAGCGGCC GCTAGCTTCT 5160
20 GTTTTGGCGG ATGAGAGAAG ATTTTCAGCC TGATACAGAT TAAATCAGAA CGCAGAAGCG 5220
GTCTGATAAA ACAGAATTG C 5241

(2) INFORMATION FOR SEQ ID NO: 3:

- 25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: RNA (genomic)
- (iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Influenza virus, 3' RNA sequence

(C) INDIVIDUAL ISOLATE: Wild Type vRNA Promotor Element

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

UCGUUUUCGU CCC

13

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

10

(A) LENGTH: 13 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(iii) HYPOTHETICAL: NO

15

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Influenza virus, vRNA 5' sequence

(C) INDIVIDUAL ISOLATE: pHL1104 vRNA Promoter Element

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

20

UCAUCUUUGU CCC

13

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

25

(A) LENGTH: 14 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

30

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Influenza virus, cRNA 3' sequence

(C) INDIVIDUAL ISOLATE: cRNA Promoter element

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

AGUAGAAACA AGGG

14

(2) INFORMATION FOR SEQ ID NO: 6:

5

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6802 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

10

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Influenza virus, RNA sequence

(C) INDIVIDUAL ISOLATE: PHL1191

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CTAGACTCTT CAAGCAAAG CAGGTAGATC TTGAAAGATG AGTCTTCTAA CCGAGGTCTGA	60
AACGTACGTT CTCTCTATCA TCCCGTCAGG CCCCCTCAAA GCCGAGATCG CACAGAGACT	120
TGAAGATGTC TTTGCAGGGA AGAACACCGA TCTTGAGGTT CTCATGGAAT GGCTAAAGAC	180
AAGACCAATC CTGTCACCTC TGACTAAGGG GATTTTAGGA TTTGTGTTCA CGCTACCGT	240
20 GCCCAGTGAG CGAGGACTGC AGCGTAGACG CTTTGTCCAA AATGCCCTTA ATGGGAACGG	300
GGATCCAAAT AACATGGACA AAGCAGTTAA ACTGTATAGG AAGCTCAAGA GGGAGATAAC	360
ATTCCATGGG GCCAAAGAAA TCTCACTCAG TTATTCTGCT GGTGCACTTG CCAGTTGTAT	420
GGGCCTCATA TACAACAGGA TGGGGGCTGT GACCACTGAA GTGGCATTG GCCTGGTATG	480
TGCAACCTGT GAACAGATTG CTGACTCCCA GCATCGGTCT CATAGGCAAA TGGTGACAAC	540
25 AACCAACCCA CTAATCAGAC ATGAGAACAG AATGGTTTTA GCCAGCACTA CAGCTAAGGC	600
TATGGAGCAA ATGGCTGGAT CGAGTGAGCA AGCAGCAGAG GCCATGGAGG TTGCTAGTCA	660
GGCTAGGCAA ATGGTGCAAG CGATGAGAAC CATTGGGACT CATCCTAGCT CCAGTGCTGG	720
TCTGAAAAAT GCTCTTCTTG AAAATTTGCA GGCCTATCAG AAACGAATGG GGGTGAGAT	780
GCAACGGTTC AAGTGATCCT CTCGCTATTG CCGCAAATAT CATTGGGATC TTGCACTTGA	840

TATTGTGGAT TCTTGATCGT CTTTTTTTCA AATGCATTTA CCGTCGCTTT AAATACGGAC 900
TGAAAGGAGG GCCTTCTACG GAAGGAGTGC CAAAGTCTAT GAGGGAAGAA TATCGAAAGG 960
AACAGCAGAG TGCTGTGGAT GCTGACGATG GTCATTTTGT CAGCATAGAG CTGGAGTAAA 1020
AAACTACCAT ATGGGGCATG TCCCAAGTAT GTTAAGCAAA ACACCTCTGAA GTTGGCAACA 1080
5 GGGATGCGGA ATGTACCAGA GAAACAAACT AGAGGCATAT TCGGCGCAAT AGCAGGTTTC 1140
ATAGAAAATG GTTGGGAGGG AATGATAGAC GGTGGGTACG GTTTCAGGCA TCAAATTCC 1200
GAGGGCACAG GACAAGCAGC AGATCTTAAA AGCACTCAAG CAGCCATCGA CCAAATCAAT 1260
GGGAAACTGA ATAGGGTAAT CGAGAAGACG AACGAGAAAT TCCATCAAAT CGAAAAGGAA 1320
TTCTCAGAAG TAGAAGGGAG AATTCAGGAC CTCGAGCATG CATCTAGAGG GCCCTATTCT 1380
10 ATAGTGTCAC CTAAATGCTA GAGCTCGCTG ATCAGCCTCG ACTGTGCCTT CTAGTTGCCA 1440
GCCATCTGTT GTTTGCCCCCT CCCCCGTGCC TTCCTTGACC CTGGAAGGTG CCACTCCCAC 1500
TGTCCTTTCC TAATAAAATG AGGAAATTGC ATCGCATTGT CTGAGTAGGT GTCATTCTAT 1560
TCTGGGGGGT GGGGTGGGGC AGGACAGCAA GGGGGAGGAT TGGGAAGACA ATAGCAGGCA 1620
TGCTGGGGAT GCGGTGGGCT CTATGGCTTC TGAGGCGGAA AGAACCAGCT GGGGCTCTAG 1680
15 GGGGTATCCC CACGCGCCCT GTAGCGGCGC ATTAAGCGCG GCGGGTGTGG TGGTTACGCG 1740
CAGCGTGACC GCTACACTTG CCAGCGCCCT AGCGCCCGCT CCTTTCGCTT TCTTCCCTTC 1800
CTTTCTCGCC ACGTTCGCCG GCTTTCCCCG TCAAGCTCTA AATCGGGGCA TCCCTTTAGG 1860
GTTCCGATTT AGTGCTTTAC GGCACCTCGA CCCCCAAAAA CTTGATTAGG GTGATGGTTC 1920
ACGTAGTGGG CCATCGCCCT GATAGACGGT TTTTCGCCCT TTGACGTTGG AGTCCACGTT 1980
20 CTTTAATAGT GGACTCTTGT TCCAAACTGG AACAACACTC AACCCTATCT CGGTCTTATT 2040
CTTTTGATTT ATAAGGGATT TGGGGATTTT GGCCTATTGG TTAAAAAATG AGCTGATTTA 2100
ACAAAAATTT AACGCGAATT AATTCTGTGG AATGTGTGTC AGTTAGGGTG TGGAAAGTCC 2160
CCAGGCTCCC CAGGCAGGCA GAAGTATGCA AAGCATGCAT CTCAATTAGT CAGCAACCAG 2220
GTGTGGAAAG TCCCCAGGCT CCCCAGCAGG CAGAAGTATG CAAAGCATGC ATCTCAATTA 2280
25 GTCAGCAACC ATAGTCCCCG CCCTAACTCC GCCCATCCCG CCCCTAACTC CGCCCAGTTC 2340
CGCCCATTCT CCGCCCCATG GCTGACTAAT TTTTTTTTATT TATGCAGAGG CCGAGGCCGC 2400
CTCTGCCTCT GAGCTATTCC AGAAGTAGTG AGGAGGCTTT TTTGGAGGCC TAGGCTTTTG 2460

	CAAAAAGCTC CCGGGAGCTT GTATATCCAT TTTCGGATCT GATCAAGAGA CAGGATGAGG	2520
	ATCGTTTTCGC ATGATTGAAC AAGATGGATT GCACGCAGGT TCTCCGGCCG CTTGGGTGGA	2580
	GAGGCTATTC GGCTATGACT GGGCACAACA GACAATCGGC TGCTCTGATG CCGCCGTGTT	2640
	CCGGCTGTCA GCGCAGGGGC GCCCGGTTCT TTTTGTCAAG ACCGACCTGT CCGGTGCCCT	2700
5	GAATGAACTG CAGGACGAGG CAGCGCGGCT ATCGTGGCTG GCCACGACGG GCGTTCCTTG	2760
	CGCAGCTGTG CTCGACGTTG TCACTGAAGC GGGAAGGGAC TGGCTGCTAT TGGGCGAAGT	2820
	GCCGGGGCAG GATCTCCTGT CATCTCACCT TGCTCCTGCC GAGAAAGTAT CCATCATGGC	2880
	TGATGCAATG CGGCGGCTGC ATACGCTTGA TCCGGCTACC TGCCCATTCG ACCACCAAGC	2940
	GAAACATCGC ATCGAGCGAG CACGTACTCG GATGGAAGCC GGTCTTGTCG ATCAGGATGA	3000
10	TCTGGACGAA GAGCATCAGG GGCTCGCGCC AGCCGAACTG TTCGCCAGGC TCAAGGCGCG	3060
	CATGCCCCGAC GGCGAGGATC TCGTCGTGAC CCATGGCGAT GCCTGCTTGC CGAATATCAT	3120
	GGTGGAAAAT GGCCGCTTTT CTGGATTCAT CGACTGTGGC CGGCTGGGTG TGGCGGACCG	3180
	CTATCAGGAC ATAGCGTTGG CTACCCGTGA TATTGCTGAA GAGCTTGGCG GCGAATGGGC	3240
	TGACCGCTTC CTCGTGCTTT ACGGTATCGC CGCTCCCGAT TCGCAGCGCA TCGCCTTCTA	3300
15	TCGCCTTCTT GACGAGTTCT TCTGAGCGGG ACTCTGGGGT TCGAAATGAC CGACCAAGCG	3360
	ACGCCCCAACC TGCCATCACG AGATTTTCGAT TCCACCGCCG CCTTCTATGA AAGGTTGGGC	3420
	TTCGGAATCG TTTTCCGGGA CGCCGGCTGG ATGATCCTCC AGCGCGGGGA TCTCATGCTG	3480
	GAGTTCTTCG CCCACCCCAA CTTGTTTATT GCAGCTTATA ATGGTTACAA ATAAAGCAAT	3540
	AGCATCACAA ATTTACAAA TAAAGCATTT TTTTCACTGC ATTCTAGTTG TGGTTTGTCC	3600
20	AAACTCATCA ATGTATCTTA TCATGTCTGT ATACCGTCGA CCTCTAGCTA GAGCTTGGCG	3660
	TAATCATGGT CATAGCTGTT TCCTGTGTGA AATTGTTATC CGCTCACAAT TCCACACAAC	3720
	ATACGAGCCG GAAGCATAAA GTGTAAAGCC TGGGGTGCCT AATGAGTGAG CTAACTCACA	3780
	TTAATTGCGT TGCCTCACT GCCCGCTTTC CAGTCGGGAA ACCTGTCGTG CCAGCTGCAT	3840
	TAATGAATCG GCCAACGCGC GGGGAGAGGC GGTTCGTA TTGGGCGCTC TTCCGCTTCC	3900
25	TCGCTCACTG ACTCGCTGCG CTCGGTCGTT CGGCTGCGGC GAGCGGTATC AGCTCACTCA	3960
	AAGGCGGTAA TACGGTTATC CACAGAATCA GGGGATAACG CAGGAAAGAA CATGTGAGCA	4020
	AAAGGCCAGC AAAAGGCCAG GAACCGTAAA AAGGCCGCGT TGCTGGCGTT TTTCCATAGG	4080

CTCCGCCCCC CTGACGAGCA TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG 4140
ACAGGACTAT AAAGATACCA GGC GTTTCCC CCTGGAAGCT CCCTCGTGCG CTCTCCTGTT 4200
CCGACCCTGC CGCTTACCGG ATACCTGTCC GCCTTTCTCC CTTCGGGAAG CGTGGCGCTT 4260
TCTCAATGCT CACGCTGTAG GTATCTCAGT TCGGTGTAGG TCGTTCGCTC CAAGCTGGGC 4320
5 TGTGTGCACG AACCCCCCGT TCAGCCCGAC CGCTGCGCCT TATCCGGTAA CTATCGTCTT 4380
GAGTCCAACC CGGTAAGACA CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT 4440
AGCAGAGCGA GGTATGTAGG CGGTGCTACA GAGTTCTTGA AGTGGTGGCC TAACTACGGC 4500
TACACTAGAA GGACAGTATT TGGTATCTGC GCTCTGCTGA AGCCAGTTAC CTTCGGAAAA 4560
AGAGTTGGTA GCTCTTGATC CGGCAAACAA ACCACCGCTG GTAGCGGTGG TTTTTTTGTT 4620
10 TGCAAGCAGC AGATTACGCG CAGAAAAAAA GGATCTCAAG AAGATCCTTT GATCTTTTCT 4680
ACGGGGTCTG ACGCTCAGTG GAACGAAAAC TCACGTTAAG GGATTTTGGT CATGAGATTA 4740
TCAAAAAGGA TCTTCACCTA GATCCTTTTA AATTAAAAAT GAAGTTTTAA ATCAATCTAA 4800
AGTATATATG AGTAAACTTG GTCTGACAGT TACCAATGCT TAATCAGTGA GGCACCTATC 4860
TCAGCGATCT GTCTATTTTCG TTCATCCATA GTTGCCCTGAC TCCCCGTCGT GTAGATAACT 4920
15 ACGATACGGG AGGGCTTACC ATCTGGCCCC AGTGCTGCAA TGATACCGCG AGACCCACGC 4980
TCACCGGCTC CAGATTTATC AGCAATAAAC CAGCCAGCCG GAAGGGCCGA GCGCAGAAGT 5040
GGTCCTGCAA CTTTATCCGC CTCCATCCAG TCTATTAATT GTTGCCGGGA AGCTAGAGTA 5100
AGTAGTTCGC CAGTTAATAG TTTGCGCAAC GTTGTGCGCA TTGCTACAGG CATCGTGGTG 5160
TCACGCTCGT CGTTTGGTAT GGCTTCATTC AGCTCCGGTT CCCAACGATC AAGGCGAGTT 5220
20 ACATGATCCC CCATGTTGTG CAAAAAAGCG GTTAGCTCCT TCGGTCCCTCC GATCGTTGTC 5280
AGAAGTAAGT TGGCCGCAGT GTTATCACTC ATGGTTATGG CAGCACTGCA TAATTCTCTT 5340
ACTGTCATGC CATCCGTAAG ATGCTTTTCT GTGACTGGTG AGTACTCAAC CAAGTCATTC 5400
TGAGAATAGT GTATGCGGCG ACCGAGTTGC TCTTGCCCGG CGTCAATACG GGATAATACC 5460
GCGCCACATA GCAGAACTTT AAAAGTGCTC ATCATTGGAA AACGTTCTTC GGGGCGAAAA 5520
25 CTCTCAAGGA TCTTACCGCT GTTGAGATCC AGTTCGATGT AACCCTCTCG TGCACCCAAC 5580
TGATCTTCAG CATCTTTTAC TTTCACCAGC GTTCTGGGT GAGCAAAAAC AGGAAGGCAA 5640
AATGCCGCAA AAAAGGGAAT AAGGGCGACA CGGAAATGTT GAATACTCAT ACTCTTCCTT 5700

TTTCAATATT ATTGAAGCAT TTATCAGGGT TATTGTCTCA TGAGCGGATA CATATTTGAA 5760
TGTATTTAGA AAAATAAACA AATAGGGGTT CCGCGCACAT TTCCCCGAAA AGTGCCACCT 5820
GACGTCGACG GATCGGGAGA TCTCCCGATC CCCTATGGTC GACTCTCAGT ACAATCTGCT 5880
CTGATGCCGC ATAGTTAAGC CAGTATCTGC TCCCTGCTTG TGTGTTGGAG GTCGCTGAGT 5940
5 AGTGCGCGAG CAAAATTTAA GCTACAACAA GGCAAGGCTT GACCGACAAT TGCATGAAGA 6000
ATCTGCTTAG GGTTAGGCGT TTTGCGCTGC TTCGCGATGT ACGGGCCAGA TATACGCGTT 6060
GACATTGATT ATTGACTAGT TATTAATAGT AATCAATTAC GGGGTCATTA GTTCATAGCC 6120
CATATATGGA GTTCCGCGTT ACATAACTTA CGGTAAATGG CCCGCCTGGC TGACCGCCCA 6180
ACGACCCCCG CCCATTGACG TCAATAATGA CGTATGTTCC CATAGTAACG CCAATAGGGA 6240
10 CTTTCCATTG ACGTCAATGG GTGGACTATT TACGGTAAAC TGCCCACTTG GCAGTACATC 6300
AAGTGTATCA TATGCCAAGT ACGCCCCCTA TTGACGTCAA TGACGGTAAA TGGCCCGCCT 6360
GGCATTATGC CCAGTACATG ACCTTATGGG ACTTTCCTAC TTGGCAGTAC ATCTACGTAT 6420
TAGTCATCGC TATTACCATG GTGATGCGGT TTTGGCAGTA CATCAATGGG CGTGGATAGC 6480
GGTTTGACTC ACGGGGATTT CCAAGTCTCC ACCCCATTGA CGTCAATGGG AGTTTGTTTT 6540
15 GGCACCAAAA TCAACGGGAC TTTCCAAAAT GTCGTAACAA CTCCGCCCCA TTGACGCAAA 6600
TGGGCGGTAG GCGGTACGG TGGGAGGTCT ATATAAGCAG AGCTCTCTGG CTAAC TAGAG 6660
AACCCACTGC TTA CTGGCTT ATCGAAATTA ATACGACTCA CTATAGGGAG ACCCAAGCTT 6720
GGTACCGAGC TCGGATCCAC TAGTAACGGC CGCCAGTGTG CTGGAATTCT GCAGATTCTT 6780
CTCTCATCCG CCAAAACAGA AG 6802

20 (2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5825 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
25 (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

30 (A) ORGANISM: Influenza virus, RNA sequence

(C) INDIVIDUAL ISOLATE: pHL1489

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

	CGATGGTCAT TTTGTCCTGA TGAGTCCGTG AGGACGAAAC ATAGAGCTGG AGTAAACTG	60
	ATGAGTCCGT GAGGACGAAA CTACCTTGTT TCTATTCGAA ATGACCGACC AAGCGACGCC	120
5	CAACCTGCCA TCACGAGATT TCGATTCCAC CGCCGCCTTC TATGAAAGGT TGGGCTTCGG	180
	AATCGTTTTT CGGGACGCCG GCTGGATGAT CCTCCAGCGC GGGGATCTCA TGCTGGAGTT	240
	CTTCGCCCCAC CCCGGGCTCG ATCCCCTCGC GAGTTGGTTC AGCTGCTGCC TGAGGCTGGA	300
	CGACCTCGCG GAGTTCTACC GGCAGTGCAA ATCCGTCGGC ATCCAGGAAA CCAGCAGCGG	360
	CTATCCGCGC ATCCATGCCC CCGAACTGCA GGAGTGGGGA GGCACGATGG CCGCTTTGGT	420
10	CCCGGATCTT TGTGAAGGAA CCTTACTTCT GTGGTGTGAC ATAATTGGAC AAACCTACCTA	480
	CAGAGATTTA AAGCTCTAAG GTAAATATAA AATTTTTAAG TGTATAATGT GTTAAACTAC	540
	TGATTCTAAT TGTTTGTGTA TTTTAGATTC CAACCTATGG AACTGATGAA TGGGAGCAGT	600
	GGTGGAATGC CTTTAATGAG GAAAACCTGT TTTGCTCAGA AGAAATGCCA TCTAGTGATG	660
	ATGAGGCTAC TGCTGACTCT CAACATTCTA CTCCTCCAAA AAAGAAGAGA AAGGTAGAAG	720
15	ACCCCAAGGA CTTTCCTTCA GAATTGCTAA GTTTTTTGAG TCATGCTGTG TTTAGTAATA	780
	GAACCTTGCT TTGCTTTGCT ATTTACACCA CAAAGGAAAA AGCTGCACTG CTATACAAGA	840
	AAATTATGGA AAAATATTCT GTAACCTTTA TAAGTAGGCA TAACAGTTAT AATCATAACA	900
	TACTGTTTTT TCTTACTCCA CACAGGCATA GAGTGTCTGC TATTAATAAC TATGCTCAAA	960
	AATTGTGTAC CTTTAGCTTT TTAATTTGTA AAGGGGTAA TAAGGAATAT TTGATGTATA	1020
20	GTGCCTTGAC TAGAGATCAT AATCAGCCAT ACCACATTG TAGAGGTTTT ACTTGCTTTA	1080
	AAAAACCTCC CACACCTCCC CCTGAACCTG AACATAAAA TGAATGCAAT TGTGTGTGTT	1140
	AACCTGTTTA TTGCAGCTTA TAATGGTTAC AAATAAAGCA ATAGCATCAC AAATTCACA	1200
	AATAAAGCAT TTTTTTCACT GCATTCTAGT TGTGGTTTGT CCAAATCAT CAATGTATCT	1260
	TATCATGTCT GGATCCCCAG GAAGCTCCTC TGTGTCCTCA TAAACCCTAA CCTCCTCTAC	1320
25	TTGAGAGGAC ATTCCAATCA TAGGCTGCCC ATCCACCCTC TGTGTCCTCC TGTTAATTAG	1380
	GTCACCTAAC AAAAAGGAAA TTGGGTAGGG GTTTTTTACA GACCGCTTTC TAAGGGTAAT	1440

	TTTAAATAT CTGGGAAGTC CCTTCCACTG CTGTGTTCCA GAAGTGTGG TAAACAGCCC	1500
	ACAAATGTCA ACAGCAGAAA CATAAAGCT GTCAGCTTTG CACAAGGGCC CAACACCCTG	1560
	CTCATCAAGA AGCACTGTGG TTGCTGTGT AGTAATGTGC AAAACAGGAG GCACATTTTC	1620
	CCCACCTGTG TAGGTTCCAA AATATCTAGT GTTTTCATTT TTA CTGAT CAGGAACCCA	1680
5	GCACTCCACT GGATAAGCAT TATCCTTATC CAAAACAGCC TTGTGGTCAG TGTTCATCTG	1740
	CTGACTGTCA ACTGTAGCAT TTTTGGGGT TACAGTTTGA GCAGGATATT TGGTCCTGTA	1800
	GTTTGCTAAC ACACCCTGCA GCTCCAAAGG TTCCCCACCA ACAGCAAAAA AATGAAAATT	1860
	TGACCCTTGA ATGGGTTTT CAGCACCATT TTCATGAGTT TTTGTGTCC CTGAATGCAA	1920
	GTTTAACATA GCAGTTACCC CAATAACCTC AGTTTTAACA GTAACAGCTT CCCACATCAA	1980
10	AATATTTCCA CAGGTAAAGT CCTCATTTAA ATTAGGCAAA GGAATTCTTG AAGACGAAAG	2040
	GGCCTCGTGA TACGCCTATT TTTATAGGTT AATGTCATGA TAATAATGGT TTCTTAGACG	2100
	TCAGGTGGCA CTTTTCGGG AAATGTGCGC GGAACCCCTA TTTGTTTATT TTTCTAAATA	2160
	CATTCAAATA TGTATCCGCT CATGAGACAA TAACCCTGAT AAATGCTTCA ATAATATTGA	2220
	AAAAGGAAGA GTATGAGTAT TCAACATTT CGTGTCGCCC TTATTCCCTT TTTGCGGCA	2280
15	TTTTGCCTTC CTGTTTTTGC TCACCCAGAA ACGCTGGTGA AAGTAAAAGA TGCTGAAGAT	2340
	CAGTTGGGTG CACGAGTGGG TTACATCGAA CTGGATCTCA ACAGCGGTAA GATCCTTGAG	2400
	AGTTTTCGCC CCGAAGAACG TTTTCCAATG ATGAGCACTT TTAAAGTTCT GCTATGTGGC	2460
	GCGGTATTAT CCCGTGTTGA CGCCGGGCAA GAGCAACTCG GTCGCCGCAT AACTATTCT	2520
	CAGAATGACT TGGTTGAGTA CTCACCAGTC ACAGAAAAGC ATCTTACGGA TGGCATGACA	2580
20	GTAAGAGAAT TATGCAGTGC TGCCATAACC ATGAGTGATA AACTGCGGC CAACTTACTT	2640
	CTGACAACGA TCGGAGGACC GAAGGAGCTA ACCGCTTTTT TGCACAACAT GGGGGATCAT	2700
	GTAACCTGCC TTGATCGTTG GGAACCGGAG CTGAATGAAG CCATACCAAA CGACGAGCGT	2760
	GACACCACGA TGCCTGCAGC AATGGCAACA ACGTTGCGCA AACTATTAAC TGGCGAACTA	2820
	CTTACTCTAG CTCCCCGGA ACAATTAATA GACTGGATGG AGGCGGATAA AGTTGCAGGA	2880
25	CCACTTCTGC GTCGGCCCT TCCGGCTGGC TGGTTTATTG CTGATAAATC TGGAGCCGGT	2940
	GAGCGTGGGT CTCGCGGTAT CATTGCAGCA CTGGGGCCAG ATGGTAAGCC CTCCCGTATC	3000
	GTAGTTATCT ACACGACGGG GAGTCAGGCA ACTATGGATG AACGAAATAG ACAGATCGCT	3060

GAGATAGGTG CCTCACTGAT TAAGCATTGG TAACTGTCAG ACCAAGTTTA CTCATATATA 3120
CTTTAGATTG ATTTAAAACT TCATTTTAA TTTAAAAGGA TCTAGGTGAA GATCCTTTTT 3180
GATAATCTCA TGACCAAAAT CCCTTAACGT GAGTTTTTCGT TCCACTGAGC GTCAGACCCC 3240
GTAGAAAAGA TCAAAGGATC TTCTTGAGAT CCTTTTTTTC TGCGCGTAAT CTGCTGCTTG 3300
5 CAAACAAAAA AACCACCGCT ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT 3360
CTTTTTCCGA AGGTAACTGG CTTCAGCAGA GCGCAGATAC CAAATACTGT CCTTCTAGTG 3420
TAGCCGTAGT TAGGCCACCA CTTCAAGAAC TCTGTAGCAC CGCCTACATA CCTCGCTCTG 3480
CTAATCCTGT TACCAGTGGC TGCTGCCAGT GGCATAAGT CGTGTCTTAC CGGGTTGGAC 3540
TCAAGACGAT AGTTACCGGA TAAGGCGCAG CGGTCGGGCT GAACGGGGGG TTCGTGCACA 3600
10 CAGCCCAGCT TGGAGCGAAC GACCTACACC GAACTGAGAT ACCTACAGCG TGAGCTATGA 3660
GAAAGCGCCA CGCTTCCCGA AGGGAGAAAG GCGGACAGGT ATCCGGTAAG CGGCAGGGTC 3720
GGAACAGGAG AGCGCACGAG GGAGCTTCCA GGGGAAACG CCTGGTATCT TTATAGTCCT 3780
GTCGGGTTTC GCCACCTCTG ACTTGAGCGT CGATTTTTGT GATGCTCGTC AGGGGGGCGG 3840
AGCCTATGGA AAAACGCCAG CAACGCGGCC TTTTACGGT TCCTGGCCTT TTGCTGGCCT 3900
15 TTTGCTCACA TGTTCTTTCC TCGGTTATCC CCTGATTCTG TGGATAACCG TATTACCGCC 3960
TTTGAGTGAG CTGATACCGC TCGCCGAGC CGAACGACCG AGCGCAGCGA GTCAGTGAGC 4020
GAGGAAGCGG AAGAGCGCCT GATGCGGTAT TTTCTCCTTA CGCATCTGTG CGGTATTTCA 4080
CACCGCATAT GGTGCACTCT CAGTACAATC TGCTCTGATG CCGCATAGTT AAGCCAGTAT 4140
ACACTCCGCT ATCGCTACGT GACTGGGTCA TGGCTGCGCC CCGACACCCG CCAACACCCG 4200
20 CTGACGCGCC CTGACGGGCT TGTCTGCTCC CGGCATCCGC TTACAGACAA GCTGTGACCG 4260
TCTCCGGGAG CTGCATGTGT CAGAGGTTTT CACCGTCATC ACCGAAACGC GCGAGGCAGC 4320
TGTGGAATGT GTGTCAGTTA GGGTGTGGAA AGTCCCCAGG CTCCCCAGCA GGCAGAAGTA 4380
TGCAAAGCAT GCATCTCAAT TAGTCAGCAA CCAGGTGTGG AAAGTCCCCA GGCTCCCCAG 4440
CAGGCAGAAG TATGCAAAGC ATGCATCTCA ATTAGTCAGC AACCATAGTC CCGCCCCTAA 4500
25 CTCCGCCCAT CCGCCCCTA ACTCCGCCCA GTTCCGCCCA TTCTCCGCC CATGGCTGAC 4560
TAATTTTTTT TATTTATGCA GAGGCCGAGG CCGCCTCGGC CTCTGAGCTA TTCCAGAAGT 4620
AGTGAGGAGG CTTTTTGGG GGCCTAGGCT TTTGCAAAAA GCTTCACGCT GCCGCAAGCA 4680

CTCAGGGCGC AAGGGCTGCT AAAGGAAGCG GAACACGTAG AAAGCCAGTC CGCAGAAACG 4740
 GTGCTGACCC CGGATGAATG TCAGCTACTG GGCTATCTGG ACAAGGGAAA ACGCAAGCGCⁿ 4800
 AAAGAGAAAAG CAGGTAGCTT GCAGTGGGCT TACATGGCGA TAGCTAGACT GGGCGGTTTT 4860
 ATGGACAGCA AGCGAACCGG AATTGCCAGC TGGGGCGCCC TCTGGTAAGG TTGGGAAGCC 4920
 5 CTGCAAAGTA AACTGGATGG CTTTCTTGCC GCCAAGGATC TGATGGCGCA GGGGATCAAG 4980
 ATCTGATCAA GAGACAGGAT GAGGATCGTT TCGCATGATT GAACAAGATG GATTGCACGC 5040
 AGGTTCTCCG GCCGCTTGGG TGGAGAGGCT ATTCCGCTAT GACTGGGCAC AACAGACAAT 5100
 CGGCTGCTCT GATGCCGCCG TGTTCGGCT GTCAGCGCAG GGGCGCCCGG TTCTTTTTGT 5160
 CAAGACCGAC CTGTCCGGTG CCCTGAATGA ACTGCAGGAC GAGGCAGCGC GGCTATCGTG 5220
 10 GCTGGCCACG ACGGGCGTTC CTTGCGCAGC TGTGCTCGAC GTTGCTACTG AAGCGGGAAG 5280
 GGACTGGCTG CTATTGGGCG AAGTGCCGGG GCAGGATCTC CTGTCATCTC ACCTTGCTCC 5340
 TGCCGAGAAA GTATCCATCA TGGCTGATGC AATGCGGCGG CTGCATACGC TTGATCCGGC 5400
 TACCTGCCCA TTCGACCACC AAGCGAAACA TCGCATCGAG CGAGCACGTA CTCGGATGGA 5460
 AGCCGGTCTT GTCGATCAGG ATGATCTGGA CGAAGAGCAT CAGGGGCTCG CGCCAGCCGA 5520
 15 ACTGTTCCGC AGGCTCAAGG CGCGCATGCC CGACGGCGAG GATCTCGTCG TGACCCATGG 5580
 CGATGCCTGC TTGCCGAATA TCATGGTGGG AAATGGCCGC TTTTCTGGAT TCATCGACTG 5640
 TGGCCGGCTG GGTGTGGCGG ACCGCTATCA GGACATAGCG TTGGCTACCC GTGATATTGC 5700
 TGAAGAGCTT GGC GGCGAAT GGGCTGACCG CTCCTCGTG CTTTACGGTA TCGCCGCTCC 5760
 CGATTCGCAG CGCATCGCCT TCTATCGCCT TCTTGACGAG TTCTTCTGAG CGGGACTCTG 5820
 20 GGGTT 5825

(2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4023 base pairs
 (B) TYPE: nucleic acid
 25 (C) STRANDEDNESS: double
 (D) TOPOLOGY: circular

 (ii) MOLECULE TYPE: DNA (genomic)

 (iii) HYPOTHETICAL: NO

 (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE: .

(A) ORGANISM: Influenza virus, RNA sequence

(C) INDIVIDUAL ISOLATE: pHL1490

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

5	CAGGTACATA TTGAAAGATG AGTCTTCTAA CCGAGGTCGA AACGTACGTT CTCTCTATCA	60
	TCCCCTCAGG CCCCCTCAAA GCCGAGATCG CACAGAGACT TGAAGATGTC TTTGCAGGGA	120
	AGAACACCGA TCTTGAGGTT CTCATGGAAT GGCTAAAGAC AAGACCAATC CTGTCACCTC	180
	TGACTAAGGG GATTTTAGGA TTTGTGTTCA CGCTCACCGT GCCCAGTGAG CGAGGACTGC	240
	AGCGTAGACG CTTTGTCCAA AATGCCCTTA ATGGGAACGG GGATCCAAAT AACATGGACA	300
10	AAGCAGTTAA ACTGTATAGG AAGCTCAAGA GGGAGATAAC ATTCCATGGG GCCAAAGAAA	360
	TCTCACTCAG TTATTCTGCT GGTGCACTTG CCAGTTGTAT GGGCCTCATA TACAACAGGA	420
	TGGGGGCTGT GACCACTGAA GTGGCATTG GCCTGGTATG TGCAACCTGT GAACAGATTG	480
	CTGACTCCCA GCATCGGTCT CATAGGCAAA TGGTGACAAC AACCAACCCA CTAATCAGAC	540
	ATGAGAACAG AATGGTTTTA GCCAGCACTA CAGCTAAGGC TATGGAGCAA ATGGCTGGAT	600
15	CGAGTGAGCA AGCAGCAGAG GCCATGGAGG TTGCTAGTCA GGCTAGGCAA ATGGTGCAAG	660
	CGATGAGAAC CATTGGGACT CATCCTAGCT CCAGTGCTGG TCTGAAAAAT GCTCTTCTTG	720
	AAAATTTGCA GGCCTATCAG AAACGAATGG GGGTGAGAT GCAACGGTTC AAGTGATCCT	780
	CTCGCTATTG CCGCAAATAT CATTGGGATC TTGCACTTGA TATTGTGGAT TCTTGATCGT	840
	CTTTTTTTTCA AATGCATTTA CCGTCGCTTT AAATACGGAC TGAAAGGAGG GCCTTCTACG	900
20	GAAGGAGTGC CAAAGTCTAT GAGGGAAGAA TATCGAAAGG AACAGCAGAG TGCTGTGGAT	960
	GCTGACGATG GTCATTTTGT CAGTATAGAG CTGGAGTAAA AAAGTACCTT GTTTCTACTA	1020
	CCTATCTCCA GGTCCAATAG GACCAGATCT AAAAGATCAC AAGCATAAAA GAGACAGGGA	1080
	GGAAAGTGAC AGGCCACAGA GAATACCTGG AAGTCATACC TGGGGAGGTG GCCCAAAAAT	1140
	GACCCCATAA CGAAAAGAAC CGGACCTCAA AGGAACAACCT GGTCGACCTG ACAACCCGGA	1200
25	GAAGTGATAA GACCGACAGG TCAATGAAAG AAAACCGCGC CCATAGCTTT CCCTCGGCCT	1260
	CAGACGTACA GGGCCCGCTA ACAGCGCGAT TTGCTGGTGA CCAATGCGA CCAGATGCTC	1320
	CACGCCCAGT CGCGTACCGT CTTATGGGA GAAAATAATA CTGTTGATGG GTGTCTGGTC	1380

AGAGACATCA AGAAATAACG CCGGAACATT AGTGCAGGCA GCTTCCACAG CAATGGCATC 1440
CTGGTCATCC AGCGGATAGT TAATGATCTA CCGCGGTAGA TCTAGCCCAC TGAACGCGGG 1500
CGGCACCTCG CTAACGGATT CACCACTCCA AGAATTGGAG CCAATCAATT CTTGCGGAGA 1560
ACTGTGAATG CGCAAACCAA CCCTTGGCAG AACATATCCA TCGCGTCCGC CATCTCCAGC 1620
5 AGCCGCACGC GCGGCATCTC GGGCCGCGTT GCTGGCGTTT TTCCATAGGC TCCGCCCCCC 1680
TGACGAGCAT CACAAAAATC GACGCTCAAG TCAGAGGTGG CGAAACCCGA CAGGACTATA 1740
AAGATACCAG GCGTTTCCCC CTGGAAGCTC CCTCGTGCGC TCTCCTGTTT CGACCCTGCC 1800
GCTTACCGGA TACCTGTCCG CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT CTCAATGCTC 1860
ACGCTGTAGG TATCTCAGTT CGGTGTAGGT CGTTCGCTCC AAGCTGGGCT GTGTGCACGA 1920
10 ACCCCCCGTT CAGCCCGACC GCTGCGCCTT ATCCGGTAAC TATCGTCTTG AGTCCAACCC 1980
GGTAAGACAC GACTTATCGC CACTGGCAGC AGCCACTGGT AACAGGATTA GCAGAGCGAG 2040
GTATGTAGGC GGTGCTACAG AGTTCTTGAA GTGGTGGCCT AACTACGGCT AACTAGAAAG 2100
GACAGTATTT GGTATCTGCG CTCTGCTGAA GCCAGTTACC TTCGGAAAAA GAGTTGGTAG 2160
CTCTTGATCC GGCAAACAAA CCACCGCTGG TAGCGGTGGT TTTTTGTTT GCAAGCAGCA 2220
15 GATTACGCGC AGAAAAAAG GATCTCAAGA AGATCCTTTG ATCTTTTCTA CGGGGTCTGA 2280
CGCTCAGTGG AACGAAACT CACGTTAAGG GATTTTGGTC ATGAGATTAT CAAAAAGGAT 2340
CTTCACCTAG ATCCTTTTAA ATTAAAAATG AAGTTTTTAA TCAATCTAAA GTATATATGA 2400
GTAAACTTGG TCTGACAGTT ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG 2460
TCTATTTCTG TCATCCATAG TTGCCTGACT CCCCGTCGTG TAGATAACTA CGATACGGGA 2520
20 GGGCTTACCA TCTGGCCCCA GTGCTGCAAT GATACCGCGA GACCCACGCT CACCGGCTCC 2580
AGATTTATCA GCAATAAACC AGCCAGCCGG AAGGGCCGAG CGCAGAAGTG GTCCTGCAAC 2640
TTTATCCGCC TCCATCCAGT CTATTAATTG TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC 2700
AGTTAATAGT TTGCGCAACG TTGTTGCCAT TGCTGCAGGC ATCGTGGTGT CACGCTCGTC 2760
GTTTGGTATG GCTTCATTCA GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC 2820
25 CATGTTGTGC AAAAAAGCGG TTAGCTCCTT CGGTCCTCCG ATCGTTGTCA GAAGTAAGTT 2880
GGCCGCAGTG TTATCACTCA TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC 2940
ATCCGTAAGA TGCTTTTCTG TGACTGGTGA GTACTCAACC AAGTCATTCT GAGAATAGTG 3000

TATGCGGCGA CCGAGTTGCT CTTGCCCCGGC GTCAACACGG GATAATACCG CGCCACATAG 3060
CAGAACTTTA AAAGTGCTCA TCATTGGAAA ACGTTCTTCG GGGCGAAAAC TCTCAAGGAT 3120
CTTACCGCTG TTGAGATCCA GTTCGATGTA ACCCACTCGT GCACCCAACT GATCTTCAGC 3180
ATCTTTTACT TTCACCAGCG TTTCTGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA 3240
5 AAAGGGAATA AGGGCGACAC GGAAATGTTG AATACTCATA CTCTTCCTTT TTCAATATTA 3300
TTGAAGCATT TATCAGGGTT ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA 3360
AAATAAACAA AAGAGTTTGT AGAAACGCAA AAAGGCCATC CGTCAGGATG GCCTTCTGCT 3420
TAATTTGATG CCTGGCAGTT TATGGCGGGC GTCCTGCCCG CCACCCTCCG GGCCGTTGCT 3480
TCGCAACGTT CAAATCCGCT CCCGGCGGAT TTGTCTTACT CAGGAGAGCG TTCACCGACA 3540
10 AACAACAGAT AAAACGAAAG GCCCAGTCTT TCGACTGAGC CTTTCGTTTT ATTTGATGCC 3600
TGGCAGTTCC CTACTCTCGC ATGGGGAGAC CCCACACTAC CATCGGCGCT ACGGCGTTTC 3660
ACTTCTGAGT TCGGCATGGG GTCAGGTGGG ACCACCGCGC TACTGCCGCC AGGCAAATTC 3720
TGTTTTATCA GACCGCTTCT GCGTTCTGAT TTAATCTGTA TCAGGCTGAA AATCTTCTCT 3780
CATCCGCCAA AACAGAAGCT AGCGGCCGCT AGCTTGGGCT GCAGGTCGAC TCTAGAGGAT 3840
15 CCCCCAAAAA AAAAAAAAAA AAAAAAAAAA GAGTCCAGAG TGGCCCCGCC GTTCCGCGCC 3900
GGGGGGGGGG GGGGGGGGGG ACACTTTCGG ACATCTGGTC GACCTCCAGC ATCGGGGGAA 3960
AAAAAAAAAA CAAAGTTTCG CCCGGAGTAC TGGTCGACCT CCGAAGTTGG GGGGGAGTAG 4020
AAA 4023

CLAIMS:

1. A segmented RNA virus characterized in that it comprises one or more segments which have been genetically modified to show improved transcription, replication and/or expression rates and which may contain non viral genes to be transferred to the host.
5
2. The virus of claim 1, wherein one or more modifications have been introduced in the noncoding region(s) and/or in the coding region(s).
3. The virus of any of claims 1 or 2, wherein at least one modified segment is derived from an original one by sequence variation(s).
- 10 4. The virus of any of claims 1 or 2, wherein at least one modified segment is an artificial addition to the set of original segments.
5. The virus of claim 4, wherein the modified segment comprises a nucleotide sequence which codes for a protein or peptide which is foreign to the original virus.
- 15 6. The virus of claim 5, wherein the foreign protein or peptide constitutes an antigen or antigen-like sequence, a T-cell epitope or related sequence.
7. The virus of any of claims 5 or 6, wherein the segment comprises repetitions of an antigen or epitope or other peptide or protein..
8. The virus of any of claims 7 to 9, wherein the antigen or epitope is derived from HIV, Herpes-Virus, Rhinovirus, CMV, papilloma viruses, Hepatitis viruses and other human viruses or Hog Cholera Virus.
20
9. A pharmaceutical preparation of a vaccine comprising the virus of any of claims 1 to 8.
10. Use of the virus of any of claims 1 to 8 for the preparation of pharmaceuticals.
25

INTERNATIONAL SEARCH REPORT

Intern. Patent Application No
PCT 95/03663

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/86 C12N7/00 C12N15/62 A61K39/145 A61K35/76

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CARRASCO, L. et al. 'Regulation of gene expression in animal viruses'; 1993, PLENUM PRESS, NEW YORK pages 107-114, GARCIA-SASTRE, A. & PALESE, P. : 'Infectious influenza viruses from cDNA-derived RNA: reverse genetics' see page 109, line 10 - line 11 see page 112, line 17 - page 113 --- -/--	1-10

☒ Further documents are listed in the continuation of box C.☐ Patent family members are listed in annex.

* Special categories of cited documents :

* 'A' document defining the general state of the art which is not considered to be of particular relevance

* 'E' earlier document but published on or after the international filing date

* 'L' document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

* 'O' document referring to an oral disclosure, use, exhibition or other means

* 'P' document published prior to the international filing date but later than the priority date claimed

* 'T' later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

* 'X' document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

* 'Y' document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

* '&' document member of the same patent family

Date of the actual completion of the international search

8 February 1996

Date of mailing of the international search report

27. 02.96

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,
Fax (+ 31-70) 340-3016

Authorized officer

Andres, S

INTERNATIONAL SEARCH REPORT

Inter. Application No.

PCI/EP 95/03663

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Quotation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	J GEN VIROL 73 (12). 3115-3124, SEONG, B. ET AL. 'NUCLEOTIDES 9 TO 11 OF THE INFLUENZA A VIRION RNA PROMOTER ARE CRUCIAL FOR ACTIVITY IN-VITRO.' see page 3117, left column, line 14 - right column, line 2 see page 3119, paragraph 2 - page 3120 see table 4 see page 3123, right column, line 8 - page 3123 ---	1-10
Y	CELL, vol. 59, no. 6, 22 December 1989 CAMBRIDGE, MA US, pages 1107-1113, LUYTJES, W. ET AL. 'Amplification, expression, and packaging of a foreign gene by influenza virus' see the whole document ---	1-10
A	JOURNAL OF VIROLOGY, vol. 67, no. 11, pages 6659-6666, LI, S. ET AL. 'Chimeric influenza virus induces neutralizing antibodies and cytotoxic T cells against human immunodeficiency virus type 1' see the whole document ---	1-10
A	J VIROL 66 (7). 1992. 4331-4338, LI, X. ET AL. 'MUTATIONAL ANALYSIS OF THE PROMOTER REQUIRED FOR FLUENZA VIRUS VIRION RNA SYNTHESIS.' cited in the application ---	
A	VIRUS RES 28 (2). 1993. 99-112, PICCONE, M. ET AL. 'MUTATIONAL ANALYSIS OF THE INFLUENZA VIRUS VRNA PROMOTER.' ---	
A	JOURNAL OF VIROLOGY, vol. 68, no. 6, pages 4092-4096, FODOR, E. ET AL. 'The influenza virus panhandle is involved in the initiation of transcription' cited in the application ---	
P,X	JOURNAL OF GENERAL VIROLOGY 76 (7). 1709-1717, July 1995 NEUMANN, G. ET AL. 'Mutational analysis of influenza virus promoter elements in vivo.' see the whole document ---	1-10
	-/--	

INTERNATIONAL SEARCH REPORT

Inter. Appl. No.
PCT/95/03663

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
T	FALL MEETING OF THE SOCIETY OF BIOLOGICAL CHEMISTRY, WUERZBURG, GERMANY, SEPTEMBER 19-21, 1994. BIOLOGICAL CHEMISTRY HOPPE-SEYLER 375 (SPEC. SUPPL. 1). 1994. S73, XP 000561852 MENKE, A. ET AL. 'Ribozyme mediated cleavage of influenza NP-vRNA in vitro and in vivo.' ---	
T	KEYSTONE SYMPOSIUM ON RIBOZYMES: BASIC SCIENCE AND THERAPEUTIC APPLICATIONS, BRECKENRIDGE, COLORADO, USA, JANUARY 15-21, 1995. JOURNAL OF CELLULAR BIOCHEMISTRY SUPPLEMENT 0 (19A). 1995. 223, MENKE, A. ET AL. 'Double ribozyme mediated cleavage of influenza A NP-vRNA.' ---	
T	J MED VIROL, (1994 APR) 42 (4) 385-95, XP 000561857 TANG, X. ET AL. 'Ribozyme mediated destruction of influenza A virus in vitro and in vivo.' -----	

